THE LEARNING PROCESS and PROGRAMMED INSTRUCTION

THE LEARNING PROCESS and

PROGRAMMED INSTRUCTION

THE LEARNING PROCESS and PROGRAMMED INSTRUCTION

EDWARD J. GREEN
Dartmouth Medical School and College

Holt, Rinehart and Winston, Inc.
New York

Library of Congress Catalog Card Number: 62-9998
23268-0112
Printed in the United States of America

TO

Elisabeth Ann
Margaret Ellen
Brian Edward
and
Andrew Thornton

Preface

Programmed instruction is the first application of laboratory techniques utilized in the study of the learning process to the practical problems of education.

The experimental psychologist has too often avoided these practical problems because he has felt that teachers have not asked meaningful questions about learning. On the other hand, the teacher has not concerned himself with the literature of experimental psychology; studies of the learning process in the white rat or the pigeon seem to bear little relation to the task facing him in the classroom. This situation has been aggravated by the experimenter's adoption of vocabularies and concept systems which, though meaningful in the laboratory, seem unrealistic and abstract to the practitioner of education. In his training, the teacher often is not exposed to recent concepts in learning theory and consequently is not afforded an opportunity to extrapolate these concepts to his practice.

I have tried to bridge in some measure this information gap between the two groups. There are techniques available for the control of behavior that can be applied

imaginatively to practical education. Programmed instruction at present employs only the simplest of these techniques. The exploitation of techniques of behavioral control is a challenge for the experimental psychologist. I hope that the present treatment of programmed instruction will show the experimenter that the problems of education are the same as those met in the laboratory —although the stakes are higher, and the ingenuity required to impose satisfactory control is greater. At the same time, I hope that the teacher contemplating the use of programmed instruction or the teacher who may only want to know more about this technique will gain from the first half of this book sufficient familiarity with the really useful concepts and techniques of control to interest him in applying them to his own classroom procedures. I feel sure that looking at behavior from the point of view of the experimentalist can provide the teacher with fresh insights that need not dehumanize his relationship with his students, but rather afford a more effective teacher-student relationship in the best sense of the term.

I am indebted to many for the contributions that provide the substantive material of this book—not only to researchers who have recently entered the area of programmed instruction, but also to those who have developed the concepts and techniques in the animal behavior laboratories. For providing me with the time necessary to accomplish the research upon which some of this book is based and for the completion of the manuscript, I am indebted to my colleagues in the Department of Psychology; the Faculty Research Committee; Dr. Arthur C. Jensen, Dean of the Faculty; and Dr. John W. Masland, Provost of Dartmouth College. The Car-

negie Corporation of New York has most generously supported our research efforts in this area. Dr. S. Marsh Tenney, Dean of the Dartmouth Medical School, has with great personal interest and understanding set the progress of programmed instruction at Dartmouth into motion. The interest and cooperation of my many colleagues in the Dartmouth Medical School, especially Drs. Phillip O. Nice, Robert E. Gosselin, and Kenneth E. Moore, though not directly contributing to the preparation of the present work, have assisted research in this area by showing me the importance of this new development in education and the need for others to be informed about it. In particular I wish to express my appreciation to my colleague, Dr. Robert J. Weiss of the Department of Psychiatry, for his help in establishing and in conducting the Dartmouth project in programmed instruction and for the privilege of working with him.

Mrs. Elaine Rameor, Mrs. Joan O'Connell, Mrs. Carla Sykes, and Miss Elizabeth H. Doyle receive my special gratitude for having converted noted and random comments into a manuscript. Professors Theodore M. Newcomb and Robert M. Gagné have provided invaluable help in their wise and objective evaluation of the manuscript. Acknowledgment is hereby given Professor B. F. Skinner, Thomas F. Gilbert, A. A. Lumsdaine, The American Psychological Association, The National Education Association, and the University of Pittsburgh Press for permission to quote from their publications.

Appreciation for the greatest and most significant assistance is reserved for my wife, Ramona Thornton Green. Without her constant encouragement and forbearance, the book could not possibly have been completed.

Despite the best efforts of my many friends and colleagues, errors unquestionably persist in the manuscript and are committed to type. These errors I shall assume as my own unique contribution to the present work.

E. J. G.

Hanover, New Hampshire
February 1962

Contents

CHAPTER ONE

Assumptions

"Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior." [1]

No one objects to the physicist's predicting the trajectory of a billiard ball. However, when a psychologist undertakes to predict the behavior of a living organism, particularly a human organism, all kinds of objections are raised. It is said that even if it were possible to predict and control human behavior, it would be ethically undesirable to do so. Besides, it is said that scientific prediction of the behavior of the human organism is not feasible because of the inherent complexity of behavior.

These questions bear only indirectly upon the educational process. The reader may legitimately wonder why we concern ourselves with these issues in a book dealing with a technique of instruction. The answer is that these new techniques are based upon a system of knowledge for which the present issues are significant. In order fully to exploit the potential usefulness of a

[1] John B. Watson, "Psychology as the Behaviorist Views It," *Psychological Review*, 1913, pp. 158–177.

machine, it is necessary to know something of its inner workings. In order to understand the inner workings of the machine that is the analytic system upon which programmed instruction is based, we must be familiar with its basic assumptions. Just as the concern of the physicist is with the prediction and control of events and objects in space and time, so must the concern of the psychologist be with the prediction and control of behavior. Similarly, just as the activity of the physicist has led to the development of weapons of unimaginable destructive power, so may the development of a science of behavior lead to similar results. The objections must be acknowledged and answered.

With regard to the ethics of a science of behavior, then, it should be noted that a weapon—or an invention that may be used as a weapon—does not of itself constitute an ethical or a moral issue. The use to which the invention is put, however, lies within the province of ethics. The invention itself has no ethical status. It may be put to uses that are good and beneficial as well as to uses that are harmful. Historically, those enterprises which have given man increased control over his environment have benefited him in the long run. The development of a science of behavior could give man increased control of his social environment, and there is no reason to assume that the development of such a science of behavior would necessarily have harmful effects in the long run. One may nostalgically look back upon the time when war was waged with relatively simple and harmless means. Certainly, we view with misgivings the present potentialities of armaments that have developed as a byproduct, or, indeed, a waste product, of the activity of scientists. Yet we would not trade penicillin for stone axes, and as far

as a science of behavior is concerned, one must examine the alternatives before deciding to abandon or reject the development of such a science.

The present state of international affairs is the result of several thousand years of application of prescientific principles of behavior. The applications of these prescientific principles have demonstrated a spectacular lack of success in preventing precisely those calamities one might fear as consequences of the application of scientific principles of behavior. Aside from the outcomes of the alternatives of a science of behavior, there is another practical reason why we should proceed with all haste to develop such a science: whether or not we move ahead in this area, others will. The group that first develops and employs techniques of effective behavior control will necessarily have the edge in any race for survival.

The basic question, then, is whether it is possible or feasible to develop a systematic body of knowledge that qualifies as a science of behavior.

Classical Concepts and the Behavioral Sciences

Certain basic problems run through all fields of scientific inquiry. These problems are involved in the basic definition of the scientific method. Some of them arise from the assumptions underlying this method; others arise from the technique of its application. One of the problems in the logic of science is the place of idealistic concepts. One of these is the concept of order. Is there intrinsic order in the universe of which we are a part? This

question cannot be answered. It is illustrative of the kind of idealistic question that has retarded scientific progress. A similar question comes from the early history of experimental psychology. One of the earliest concerns of psychologists was the measurement of the sensory capacities of the human organism. Psychologists were concerned, for example, with the lower limit of hearing, the minimal acoustic stimulus necessary for auditory sensation. In classical psychophysics, the auditory threshold was regarded as an absolute, independently existing phenomenon. We now know the auditory threshold to be a statistical concept, a mean numerical value computed from a series of observations. Depending upon the way in which those observations are made, the value of the threshold will vary upward or downward. That the auditory threshold was regarded as absolute is indicated by the fact that psychophysical literature is replete with references to *error*. There are investigations of the error of anticipation, the error of perseveration, of stimulus errors. Any departures from the supposedly absolute threshold were regarded as errors. It was only relatively recently that the statistical nature of the threshold was recognized.

These classical idealistic ways of thinking have hindered the development of a science of behavior to the same extent that the concepts of absolute time and space in Newtonian physics retarded the development of relativity. The concepts themselves may be indispensable and logical steps in the development of science. But their retention and the emotional investment which scientists have in these concepts rigidifies the formulation of problems, so that new ways of looking at nature are made

difficult. In a real sense, the question of whether or not the universe is ordered is meaningless. Order, should it exist, is known only through the interaction of knowing organisms with that universe. One cannot know the universe, ordered or not, in the absence of such interactions. Scientific enterprise makes certain limited assumptions about order and includes in a growing body of knowledge observations that are congruent with such assumptions. In a way, then, science consists of the imposition of order upon the universe.

Physics is the result of the behavior of physicists. The initial assumptions that a scientist makes in studying nature will necessarily determine the kind of organization he derives. The prevailing philosophy of man in the late nineteenth century dictated to a large extent the methodologies employed in early psychological experiments. The growth of science seems to consist at times of fitting successive orders upon the universe with varying degrees of success. Concepts and principles that are less successful in describing nature are abandoned in favor of more successful ones. This is a slow process. The lesson we can learn from this is that our concepts and principles should be developed inductively from observation. As we proceed, we will also employ deductive reasoning whenever generalizations that permit deduction are reached. But basic principles must always be grounded in direct observation.

Another underlying assumption in science has to do with the notion of causality. Does A inevitably lead to *B*? This question is basic, however it may be phrased. If one observes upon one occasion that B is a consequence of A but that on no other occasion is this ever again true, the

entire scientific endeavor would collapse. The concept of causality has probably generated more argument than any other concept in the philosophy of science, and the nature of causality is a particularly cogent problem for the study of behavior.

The question of causality is so basic as to have engendered incompatible schools of psychology. Let us look at a concrete experimental situation. Stimulating a particular portion of the motor cortex of the cat by means of implanted electrodes results in the flexion of the right forepaw. One might say that electrical stimulation causes flexion of the right forepaw. However, without changing the location of the electrode and without changing the current applied, but by changing the orientation of the cat's head, the left forepaw is activated by electrical stimulation. The electrical stimulus applied to the identical portion of the cortex is the same in both instances. The stimulus input from the animal's own orientation altered its reaction. What, then, is the meaning of a statement attributing the cause of right forepaw flexion to the current from the electrode? The total stimulus situation determines the reaction. The action of the experimenter, the operation he performs upon the subject organism, is but one variable imbedded in a complex stimulus matrix. The meaning of causality intrudes into the study of behavior perhaps more, and certainly not less, than it does in other scientific disciplines.

It is appropriate now to turn to assumptions underlying scientific analyses peculiar to the behavioral sciences. The questions of order and causality are common to all scientific inquiry, but certain questions are specific to the study of behavior.

Specific Problems in the Study of Behavior

It is often argued that the behavioristic approach to the study of human behavior ignores the basic stuff of psychology—that behavior is but a surface manifestation of the real phenomena of experience. No one can deny the fact of private experience, but, on the other hand, no one has access to information concerning underlying processes of private experience except through behavior itself. One cannot directly know the private events or experiences of another organism. We may make inferences concerning such private events on the basis of that organism's or that person's behavior and the circumstances in which his behavior occurs. The only private events with which we may have direct contact are our own. If we were to be satisfied with a science of one individual then a strictly introspective position might be defensible. But because the rules of science require a concern with the repeatable, the verifiable, and the communicable, the basic data must be observables whose existence all qualified persons can agree upon. Even the most ardent mentalist must concede that what he actually observes in another is his behavior. It is only through the most confused kind of inference that one might ever assume that he could observe in others those exact and private events with which he is concerned.

To the layman, the problem of the private experience of others is not difficult. One can put aside one's professional sophistication and empathize with the layman quite easily. When I have not eaten for a time, I can observe changes in my own behavior that others can

also observe. For example, I am likely to eat when food is available. I am more likely to do things that obtain food. I can also, incidentally, observe certain reactions of mine that others cannot observe. I may label these various reactions *feelings* of hunger, or *desire* for food. I am aware, through repeated experiences, that these reactions are reliable, and they are as real to me as a table or a hot object. No one else, however, has direct knowledge of these reactions. I may speak of them, I may describe them. But, to the external observer, my speaking of them constitutes more overt behavior on my part. The observer does not observe my feelings of hunger; he observes my verbal statements about my feelings of hunger. Given my extensive private experience, however, I may form interpretations when I now observe another hungry individual. If I have some knowledge of his food deprivation, if I observe that he eats voraciously when permitted and that he does things that will obtain food for him, then I may also assume that he experiences private reactions just as I have experienced private reactions under similar circumstances. However, this assumption is unsupported by public evidence and is irrelevant. The nature of another individual's private experiences constitutes what Bridgman calls a pseudoproblem (Bridgman, 1927). That another person may, and probably does, have private experiences comparable to yours or mine is most reasonable, but it is unverifiable.

The second special problem in the study of behavior is that, unlike the relation of, say, the neutrino to the physicist, the object under investigation is the same as the behavior of the investigator, within certain broad limits. The investigator may carry with him a set of interpretations of his own behavior that cripple his objec-

tivity. The experimenter may describe the behavior of the subject verbally, but the subject may also verbally describe the experimenter, the experiment, and the relation between the experiment and his own behavior. The subject's verbal analysis may make of the experiment something other than that which the experimenter intended. This creates problems of control for the experimenter.

Control

We recognize two types of control in scientific investigation. First, statistical control, where the effects of extraneous variables are considered to vary randomly and to cancel each other. Second, physical control, where experimental manipulations are employed to reduce extraneous stimulation. Clearly, physical control is preferable. As we shall see, our concern in studying the learning process and in studying the future of programmed instruction disposes us toward physical rather than statistical control.

Considerable literature exists concerning the so-called higher mental processes and, in particular, concerning the problem of *set*. Historically, set has been the preparatory state or condition of an organism known to have a differential effect upon the subsequent consummatory activity of that organism. Our treatment of set does not regard it as essentially different from other types of physical control. In examining the literature on set, one finds that experiments dealing with those processes almost invariably use as an independent variable either an explicit or an implicit difference in instructions

to the subject. Rather than concern ourselves with the so-called higher mental process called set, we will identify our problem as the effect of differential instructions.

Historically, verbal behavior in itself has been a holdout for the dualistic conception of man in that language has been held to be an exclusively human function. The present definition of verbal behavior, however, is that it includes all behavior whose reinforcement is mediated by another organism (Skinner, 1957). Verbal behavior is not the exclusive property of the human species; the behavior of any organism can be controlled by verbal contingencies. It is not necessary to introduce special assumptions of inaccessable processes or other mysteries into an analysis of the effects of instructions. Instructions are either verbal in the sense of having been produced by the usual means of spoken or written communication, or nonverbal, as when produced by a more implicit kind of communication where verbal manipulations are more subtle, or where an individual eventually and inevitably instructs himself. When the experimenter does not explicitly instruct his subject, self-instruction will vary more than when he does. It is practically impossible in human experimentation to avoid self-instruction, even when the experimenter deliberately refrains from instructing the subject. The subject talks to himself. Oddly enough, when one compares the behavior of verbalizing human organisms with the behavior of organisms such as pigeons and rats under otherwise identical conditions of instruction, one finds that the behaviors of these differing species are essentially alike. One either draws the conclusion that infrahuman organisms instruct themselves or that the instructions humans give themselves do not have any special effects.

Our position is that verbal behavior is simply another class of behavior occurring in the organism, and that it is just as much a lawful effect of the controlling environment as other behavior under observation.

The function of instructions is directly analogous to the physical restraint of an organism as imposed by the limits of an experimental chamber. Instructions take advantage of previous verbal training in that they limit the range of potential behavior by exploiting such conditioning history as a given organism possesses. Let us examine the range of effective control one can achieve by instructions. At one extreme, one might literally tell a subject exactly what to do in an experiment. His range of behaviors can thereby be reduced to such an extent that there is no experiment. This only shows that verbal control has been established, and that one can direct the actions of people with verbal control. At the other extreme, to assume, as do Rogerians, that refusal to control by instruction means that the individual's range of behavior is completely unrestricted is unrealistic; the situation itself limits the possible actions of the individual. The subject reacts to subtle properties of the stimulus situation as he generalizes to them from prior conditions of control. The action of the therapist or experimenter vis-à-vis the client or subject cannot be totally free from communication, because communication is not limited to that which the layman considers to be language. It is necessary, therefore, to recognize the limited function of language and to use it deliberately. The essence of the argument is that control of behavior which is exercised by verbal instruction is essentially the same as control exercised by other means of physical restraint. By taking this point of view in the present analysis, we

can directly attack the problems of education from the standpoint of the experimental study of behavior. Formal education has as its central objective the modification of the verbal repertory of the student, and its principal technique is the controlled application of instructions. Instruction is synonomous with education.

Many persons may reject this notion of the control of behavior, not only with respect to verbal instruction, but also with regard to the more general problem of control, because the concept of control carries with it the connotation of *aversive* control. We have a strong tradition of freedom in our Western civilization; any infringement of freedom results in aggressive reaction, political and otherwise. If one person attempts to establish control over another, the second individual is likely to regard it as an infringement of his freedom and will engage in aggressive countercontrol. Parallel with the development of political philosophies of individual dignity and freedom has been the development of a concept of man that places upon him a peculiar responsibility of action. We say that man possesses free will, that he is capable of assessing the effects of his actions, and that he is capable of choice. Apart from the legal and political implications of this view of man, its implications for a science of behavior are devastating. If the human organism is capable of such independent action, then human behavior can never be subject to precise experimental control. At any time, the subject can step out of the framework of controlling variables and do something completely inconsistent with the effects of those variables. Behavior would necessarily be capricious in the extreme.

The only possible working hypothesis for a science

of behavior is a deterministic one. It can be shown that a concept of free will is not central to the existence of our political institutions. It is obvious that such a philosophy is incompatible with the development of a science of behavior. It is equally obvious that people do engage in actions which they describe as resulting from free choice. How can this be reconciled with the position taken here with regard to the control of human behavior?

The control that may be derived from verbal instruction does vary from individual to individual. The communality of verbal training or experience within a culture permits the reliable achievement of a certain minimal control through a given set of verbal directions. For example, one can reasonably expect that simple directions will be followed if the subject of the experiment has learned to speak the language the experimenter uses. The general verbal control exercised by the community at large goes far beyond the mere establishment of a specific language system in an individual's behavioral repertory. Gross ethical systems are also imposed by the group through verbal instruction. The child is taught that there are certain limits to culturally sanctioned behavior. The folklore of the particular culture in which the individual receives his verbal training imposes characteristic modes of problem solving, etc. However, such limits are only of a gross type; the more specific verbal training that the individual receives among his close associates, especially from parents, has a more idiosyncratic effect on him. That one cannot anticipate the unique verbal history of a given individual introduces considerable uncertainty in the predictions an experimenter may make when his subjects are instructed in a certain manner. It has been pointed out in defense of

free will that although one may advise a person to do a given thing with his right hand, he may, in fact, do something quite different. This is especially true if the individual feels that someone is imposing control upon him. The subject may strike the would-be controller instead of following his directions, which is certainly not the outcome that the controller intended. The facts of the argument are true, but the argument itself has a serious logical flaw. To say that individuals will react aggressively to verbal control implies that under certain conditions the aggressive response is a lawful consequence of controlling variables. One cannot predict an aggressive reaction if the individual is free to respond. We do not disagree with the possibility that the subject will act contrary to instruction, for he often does. This does not constitute proof of free will, but rather indicates that the individual making the prediction is not cognizant of a sufficient number of variables. He is not aware of the idiosyncratic verbal history of the individual when he makes his prediction. One need not be an experimental psychologist to realize that a prediction of behavior addressed to an individual in our culture is often reversed. In an early study, Hull showed that certain individuals, particularly delinquent girls, when given suggestions to fall forward would not follow these suggestions but would more reliably fall backwards, and vice versa (Hull, 1933). This example affords some understanding of the kinds of contingencies that have established a response to direction.

This defense of free will resolves itself into no more than a statement of our ignorance of controlling conditions. As such, it is not a compelling argument for rejecting a deterministic point of view in the analysis of behavior.

That verbal instruction is a physical event must be accepted. One can analyze the wave form of vocal speech in the atmosphere. The acoustic properties of speech are well known. Printed speech is indisputably physical; one can even photograph or otherwise record nonvocal communication and subject it to analysis. All these forms of speech may be reduced to physical measurement. The significant aspect of language, however, is not the physical event itself but the meaning it conveys. *Meaning* is another term whose definition is in dispute. Let it suffice here to say that we shall regard meaning as the relation between the physical event of speech and its effect upon the behavior of a given individual. The meaning of a word, the meaning of a phrase, the meaning of a sign, is often particular to a given individual. The meaning of a word for one individual is not necessarily the meaning of the same word for another individual. The associations established for one person may be different from the associations for another. Again, this justifies no particular mystery but points up our ignorance of the controlling history of a particular individual. The controlling history is more or less accessible. To the extent that we ignore idiosyncratic biographies, to that extent do we allow greater uncertainty in our predictions. The amount of uncertainty to be allowed depends upon one's experimental tolerance.

Selecting a Suitable Level of Analysis

A final problem, then, is the degree of generality we aim to achieve in the summary descriptive statements that form the body of our science. On the one hand, we are limited by the extremely special circumstances surround-

ing each behavioral event. Each event is physically unique in that the particular time-space coordinates and the characteristics of physical objects within them can never recur. A precise description of this totality of variables acting upon an organism at a given instant would comprise a very special kind of summary statement. It would be a summary statement about events that cannot recur, and which would therefore be of interest only to historians. Such a statement would be of little value to scientists. At the other extreme, one may be tempted to overgeneralize. The statement, "things change," is undoubtedly a correct summary statement. It is also useless as a specific prediction. A satisfactory level of analysis, for either description or theory, cannot be determined by simple reliance upon operational definitions or upon sweeping generalities. Other criteria must be sought.

Much has been made of the desirability of simplicity in theory construction. Simplicity is not the ultimate objective either, for there are pitfalls in an uncritical adoption of the principle of parsimony. Occam's razor cannot be applied indiscriminately. It is possible to predict the bending of light rays in space around an object in terms of either Euclidian or non-Euclidian geometries. If simplicity is the ultimate criterion for theory, one should favor Euclidian geometry strictly in terms of the mathematics involved. But the adoption of Euclidian geometry would necessitate the introduction of additional complicating assumptions to account for observed distortions. It is not necessary to introduce such additional assumptions with non-Euclidian geometry. "Simplicity" must be interpreted in a broad sense; it must be defined not only in terms of the basic logics of the theoretical struc-

ture, but also in terms of the rules of application governing the relation of theoretical structure to the physical world. Procrustean assumptions designed to twist Euclidian geometry to fit the facts are neither simple nor elegant. Simplicity sometimes dictates the adoption of more complex mathematics, but mathematics that per se handle the phenomena to be explained. A so-called common sense explanation of the behavior of a child in the classroom may be appealing from the standpoint of simplicity. Attempting to account for learning in such terms, however, creates more problems than it solves. Concepts derived from the experimental analysis of behavior may not have the apparent simplicity of explanatory fictions drawn from our folklore, but the simplicity of their relation to reality is infinitely greater.

In determining what is to regulate the selection of a suitable level of analysis, several traditional principles have been considered and rejected individually, and yet their rejection must be provisional. One certainly does not deliberately adopt complexity as the criterion for a suitable level of analysis nor reject a suitable degree of generality of the summary statement. The desire for simplicity is reasonable; the danger lies in too narrow a definition of simplicity. Also, although completeness in the statement of conditions existing at the time of observation is desirable from the standpoint of scientific analysis, one must steer clear of extremes. The law of parsimony does not, nor is it intended to, provide clear-cut direction. One thing is clear: the practical rules governing experimental analysis in other disciplines apply here. The same requirements for primary data exist in the study of behavior, in physics, or in biology. There is no excuse for slipping into a nonphysical level of discourse. The subject matter of

psychological analysis must ultimately be reducible to physical measurement. In the final analysis, the level at which one works and the degree of complexity or precision, the degree of generality with which one describes and manipulates the subject matter, is determined by practical considerations. The nature of the thing investigated determines the mode of investigation. It has been proposed that the level of analysis at which one works is appropriate if lawful relationships, or relationships that can be translated into smooth curves, grow from the manipulation of the independent variables. If one finds lawful relationships at a given level of analysis, then that level of analysis is appropriate in itself and requires no justification in terms of more molecular processes. This proposal opens the door for the analysis of complex behaviors not normally open to experimental study. The solution to the problem of determining a suitable level of analysis is pragmatic.

That the scientist's behavior is controlled by his subject matter necessarily makes the scientific enterprise an empirical activity. No matter how one may rationalize the way in which he goes about studying his environment, his interaction with the physical environment is always dictated by and bounded by that physical environment in a very nontrivial sense. The activity of scientists is behavior. As behavior, it is modified by the same kinds of influences that modify the behavior of other individuals. The physical environment mediates the consequences of action in scientific investigation just as it does for the small child learning to walk. Appropriate behavior is followed by consequences which further modify that behavior. Inappropriate behavior is followed by other consequences. The result, in both the case of the

scientist and of the child learning to walk, is increased control over his environment, achieved through his own efforts.

The Revolution in Man's View of Himself

The period between the eighteenth and twentieth centuries has been a period of revolution. Perhaps none of the political revolutions that took place during that time is as important for the future history of mankind as the philosophical revolution that occurred within the same period. This revolution was the change in the concept of man in relation to his universe. The origins of this revolution may be traced to Aristotle, for it was he who made explicit the contradictions in the ancient concept of man. By so doing, he posed a problem that might otherwise have gone unrecognized for some time. The Thomistic interpretation of Aristotle's writings, by sharpening the basic contradictions, sowed the seeds of destruction of the particular point of view they were supposed to protect or defend. The real break with the ancient concept came with the emergence of the mechanistic philosophies of Cabanis and La Mettrie in the wake of the French Revolution. It remained, however, for two developments at the turn of the twentieth century to initiate effective elaboration of the new concept. One of these was the writings of Sigmund Freud, whose significant contribution was the exposition of the idea that all behavior is determined. The sharpness of this concept as presented by Freud is somewhat lost in his confused clinging to a basic dualism, but he must be credited with having had the essential idea. The second

development was the founding of behaviorism by John B. Watson. The behavioristic school provided us with the experimental techniques necessary to exploit a scientific concept of man.

The present view of man's behavior as totally determined by his environment is by no means universally accepted even today. Many feel that the integrity of man is challenged by such a concept. This reaction is directly analogous to the resistance made to the notion that the earth is not the center of the solar system when it was first proposed by Copernicus. If the importance of man in the universe hinges upon his ability to reason, however that ability may ultimately be analyzed, then man's importance is not lessened by his having reasoned with regard to his own reason. Rather, as one writer has remarked (Skinner, 1955), it constitutes another tribute to this unique capacity. The dignity of man is not well served by emotional attachments to ignorance.

The radical view of behavioral determinism caused little concern as long as it was applied solely to laboratory studies of the behavior of rats, or even of college sophomores. But recently another technological development has taken place, which inevitably thrusts this radical point of view and its consequences for human action directly into the midst of human affairs. This development is the subject of the present analysis: the application of behavioral principles to problems in education. The development of programmed instruction promises to alter drastically the nature of our educational institutions and thereby even the nature of the society within which these institutions exist. The potential import of the development of the teaching machine has been likened to that of the development of the printing press.

Although this analogy may be somewhat overdrawn, it is certainly true that the broad diffusion of otherwise inaccessible material to school systems is bound to have widespread effects. It is simply not possible to accept the technology resulting from advances in behavioral science on the one hand and, on the other, to deny the validity of the basic concepts upon which the technology was based. The existence of a more effective technique of instruction within a cultural framework that does not admit the basic concepts of the science which developed this technique constitutes a contradiction sharper than any that have gone before. A stable state of affairs cannot exist when advanced technologies are employed to perpetuate prescientific knowledge. If we attempt to apply advanced technology to this end, we will fail. We can succeed only by exploiting its concepts and techniques to the fullest.

In summary, we regard human behavior as a determined monistic phenomenon. As such, it is accessible to scientific investigation. Our objective is the analysis of the learning process in order that we may employ more effective techniques of education. The techniques we employ, the questions we ask of nature, the behavior of students in a classroom, are continuous with the technique and physical nature of all scientific endeavor. There is a difference in complexity between the behavior of the laboratory animal and that of the student solving a problem in differential equations. There is no essential difference between two organisms in the processes by which their behaviors are established.

CHAPTER TWO

Definitions

The behavior of an organism consists of a set of continuously changing, interrelated actions. Behavior is not composed of neat analytic segments, but rather of undifferentiated flux. Regularities present themselves from time to time in poorly defined groupings; the identification of determining variables and the relationships between such broadly defined behaviors at the gross observational level is a challenge from which all but the most optimistic must shrink.

To bring order out of such a chaotic stream of events requires certain violence to the subject of observation in the form of simplification and perhaps even oversimplification. One must begin with limited areas of study and systematically develop the analysis to encompass ever larger portions of behavior. The first necessary simplification is thus to impose defining limits on the range of observation. Behavioral flux may be said to be made up of chains of responses. It is worth while to examine critically the concept of the *response*.

The Response and Response Class

Any definition of response is artificial. It is imposed by the observer upon behavior. The sharply defined response does not per se emerge as a natural aspect of behavior. The physical environment defines the response and the experimenter only sharpens this definition.

To enlarge somewhat, each instance of behavior is unique in that the precise physical coordinates existing at one time have changed before the next instance takes place. Behavior is time-ordered; even if there were no other differences between two response instances, they would of necessity differ because they had taken place at different times. Other, nontrivial, effects are also manifest. One variable that controlls behavior is behavior itself. The organism that has made a response a second time differs from the organism that made that response for the first time because the physical consequences of action alter the probabilities of further action by that organism. It may be that behavioral change is irreversible, if for no other reason than that the changes in the environment in which behavior takes place are irreversible.

If behavior is unique, then what possibility is there of establishing lawful relationships between behavior and its determinants? The answer lies in the degree to which analytic procedures resolve response units. One must steer a course between the Scylla of viewing poetically the undifferentiated raw stream of behavior and the Charybdis of fragmenting behavior into nonrecurring unique units. The fact that the choice must be made illustrates one of the problems of establishing a causal

law (Frank, 1957). The resolution of this particular problem for psychology is derived from the concept of the *response class*. Groups of response instances share common properties, such as their common existence as a function of some independent variable. Stated another way, the environment, in interaction with the organism, exhibits certain consistencies to which an adaptive organism can respond. The consequent accommodation of the organism in turn produces consistent effects upon the environment. A response class is defined as composed of those behaviors which are controlled by a common environmental operation upon the organism. Although individual response instances remain unique in many other respects, they share defining properties which relate them as a class to specifiable controlling variables. Whether the response to be studied is a nerve impulse, the blink of an eye, the operation of a lever by a rat, or the recitation of a poem, the situation is the same.

The experimenter must decide what is counted as a response. A "response" must be above the noise level of the measuring instrument, whether the instrument be an electronic recording system, an infrared film, or the human eye. This may seem obvious and trivial, and the establishment of response boundaries is sometimes so natural that they are established implicitly. Rather than evidence of triviality, this circumstance is cause for concern. Where the decision to establish boundaries is not explicitly made, the consequences of the decision may be overlooked in the over-all design of the experiment, and particularly in interpretations drawn from the data. As we shall see, the consequences of improper definition of the response class can be serious in the labora-

tory. The seriousness is amplified in the less well controlled environment of the classroom.

In a conditioning experiment, it is essential that whatever is counted as a response is also subjected to differential experimental treatment. The implications of violations of this injunction must be considered. The topology of specific behavior involves a set of effectors with a physical mass different from that of other effectors. Less energy is required to blink an eye than is required to press a treadle with the foot. Secondly, the differential treatment of instances of the selected response produces systematic changes within that class. Suppose that a recording device set to measure changes in a response class that included events under the control of the independent variable also measured behavior not under this control. Variability would be introduced by factors not under study. Similarly, if the measuring device did not record *all* behaviors in the class defined by operations of the independent variable, observation would be most incomplete. Either state of affairs decreases the value of the experiment. In an experiment where, for example, the rate of lever press by a white rat is studied as a function of the degree of prior food deprivation, what might be regarded as a minor apparatus failure can jeopardize the entire experiment. The recording counter must not be activated by a depression of the lever that is more or less energetic than is required to trigger the food magazine. One cannot overemphasize the necessity for precision in these matters. It is not enough to impose gross conditions of control upon an organism. The subtle behavior-environment contingencies are all-important at *any* level of analysis.

The Stimulus

A second basic concept in the study of behavior calls for elucidation; this is the concept of the *stimulus*. The problem of the definition of the stimulus has been recently surveyed (Gibson, 1960). We have been provided with an excellent survey of the variances of usage of the term and have been shown that these variances comprise a considerable problem, not only in elegance of definition, but also in confusions in the formulation of experimental questions. Gibson proposes that a worth-while enterprise would be the establishment of a branch of physics which he called ecological physics. The purpose of ecological physics would be to develop an analytic geometry of physical objects as projected upon the sensory mosaic of an organism.

In his survey of the different usages of the term "stimulus," Gibson points to the various historical usages of the term, from the older physiological concept of the stimulus as a physical energy or, as a ratio of physical energies that activate sensory organs, to more complex concepts of stimuli as they pertain to higher order processes, to learning, to social processes, and so on. Among other observations he notes that the concreteness, the preciseness, of the physical definition or description of stimulus recede as the process under investigation becomes more complex. Gibson insists it is necessary to define a stimulus in physical terms independently of the behavior of any organism. Others concur in the desirability of independent definitions of stimuli—definitions that are independent of the behavior of an organism. Yet it is not readily apparent how such an independent defini-

tion is to come about. Let us examine some general definitions of stimulus as they have been proposed.

A stimulus has been variously defined as "a part or change in part of the environment," as "the impingement of physical energy upon a receptor," or as "that which produces a change in the behavior of an organism." These definitions of a stimulus do not seem to be contradictory; indeed, they are not necessarily contradictory. If an exception to a statement is found, then that statement is not logically true. Such a statement must be limited or corrected. It will be seen that these three definitions are not equally satisfactory. Consider the eruption of Mount Vesuvius. It is safe to presume this to be a stimulus for those natives of Naples who are physically present at the scene and who have normal vision and hearing. However, it is unlikely to exert any immediate control over the behavior of a bushman having his lunch in Australia at the same time. Yet the eruption is a "change in a part of the physical environment." Clearly, the remoteness of the part of the physical environment is crucial to the breakdown of the definition in this case.

It can be shown that remoteness does not always preclude stimulation. The behavior of the same bushman will be affected by a solar eclipse. His reaction to the solar eclipse indicates that he has been stimulated, although the events producing the stimulation occur hundreds of thousands of miles away.

It may easily be established that the entire sensory system of this same bushman is continually bathed in electromagnetic radiation generated by radio stations in Melbourne, Sydney, and elsewhere. Furthermore, it is established that he is constantly bombarded by cosmic

radiation from outer space and by atomic radiation arising from local sources; yet, without some prosthetic sensing device, he makes no differential response to these events whatsoever. The second definition is thus seen to be wanting.

To define a stimulus as that which produces a change in the behavior of an organism seems trivially correct. But introducing the relation of the stimulus to behavior is a necessary qualification of the definition. Therefore, a distinction must be made between the physical event as a *potential stimulus* or *stimulus object* and its functional employment as an *effective stimulus*. The potential stimulus is something to which the experimenter is responding at the moment, but to which the subject organism has not responded. The effective stimulus is that to which the subject organism has also come to respond. Occasionally, and, unfortunately, a subject may respond to what are for him effective stimuli, but which are only potential stimuli for the experimenter.

Once it has been shown that a particular event is a stimulus for an organism, further questions may be asked about its effects. One such question has to do with the environmental context in which a stimulus is embedded. That the effect of the stimulus is dependent upon other variables also present is not debatable. All experimental situations are special. The behavior occurring in them is determined by *all* the parameters of the situation. To state that a particular stimulus is solely responsible for an instance of behavior is naive in the extreme. At the opposite extreme, it has been argued that the total stimulus complex is such that the organism cannot be brought under the control of discrete segments of it. To state

that a physical field is composed of many effective variables is not to deny the operation of individual variables within the field.[2]

In the present treatment, the term *field* explicitly refers to the complex of physical events comprising the environmental matrix within which behavior occurs and of which behavior itself forms a part. The term *field* as it is used here is not to be confused with "cognitive" or phenomenological fields. To recapitulate, the environment is composed of a large number of potential stimuli, some of which become effective in controlling behavior because of the interaction of the organism with its environment.

All that we might say about stimuli extrinsic to the organism applies equally to stimuli that are intrinsic. We must examine operations where the existence of specifiable stimuli can at best be inferred. Functional relationships may be legitimately established between molar operations, such as witholding food for a length of time and the resulting changes in behavior, without concern for the specific effective stimuli operating upon the organism as a result of this operation (Skinner, 1938). That a given instance of behavior may be produced by unspecified, or, perhaps, even unspecifiable stimuli, and that lawful relationships may nevertheless be derived has many useful implications. For a given action, the crucial stimulus, if one may conceive of such an event,

[2] Many presentations of psychological "field theorists" are confused on this point. An analysis of psychological field theory as it relates to learning theory appears in the chapter by W. K. Estes in W. K. Estes *et al.*, *Modern Learning Theory*, New York: Appleton-Century-Crofts, Inc., 1954.

may exist within the organism. The stimulus may be in the external environment. Its location is of no consequence.

When one speaks of the organism and its environment, one should not lose sight of the fact that the organism is coextensive with its environment. In early discussions one senses the notion, never explicitly stated, that the organism is a sort of pebble in a bucket of water. There is the organism, the individual, and all around him is his environment. A more accurate analogy would be to regard the organism as a sponge in the bucket of water, for the environment pervades the organism. As we have seen, the definition of the stimulus itself as it exists in the environment is predicated upon an action of the organism to it. Although the usual distinctions can be made with respect to the location of receptors in dividing the environment into external and internal components, there is no other good reason for making a division in such terms. The organism has been likened to a torus or a doughnut, in that what is normally conceived of as environment exists on the outside of the doughnut and is responded to by appropriate distance receptors and tactile receptors. The alimentary canal is topologically represented by the hole in the doughnut and the doughnut proper, the dough, constitutes the organism. Stimuli arising within the dough operate upon appropriate interoceptors. Movements of the doughnut, accelerations in space, stimulate the proprioceptors. The essential point is that the dermal boundary does not constitute an important dividing line separating the organism from its environment. Events arising within the skin are as crucial in determining the action of the organism as events existing outside of the skin.

Stimuli, therefore, which through their microscopic and/or internal locus have been and are necessarily private to the organism, have no special metaphysical status and do not constitute excuses for the introduction of paraphysical or mentalistic notions. The only defensible working hypothesis necessarily assumes that behavioral relationships with such stimuli are essentially the same as behavioral relationships related to external, macroscopically manipulable stimuli.

To return to Gibson's argument, we make several distinctions with regard to the concept of stimulus. Stimulus events may, through appropriate conditioning procedures, be established in various functional relationships to the behavior of organisms. The functional relationship that is established for a given organism at a given time, depends upon the particular reflex contingencies developed by the experimenter or by the more general controlling environment. Stimulus and response are necessarily codefined as they are determined within this relationship with one another (Skinner, 1935). In any case, it is clear that there are physical events in the environment which may not be effective stimuli for a given organism at a specific time. Such events may nonetheless be stimulus objects or potential stimuli; when it is established that they do, in fact, control or determine in some way the behavior of an experimental organism, then they become effective stimuli. One difficulty with the Gibsonian position regarding the independent physical definition of a stimulus is that it ignores the fact that the ecological physicist who develops the *n*-dimensional analytic geometries and the psychologist who directs the ecological physicist so to develop *n*-dimensional analytic geometries, have themselves responded to the physical

objects in question. Because both psychologists and ecological physicists are organisms, subject to the same laws of behavior as other organisms, Gibson has not escaped the circularity in the definition of the stimulus he seems to have hoped for. One does escape this circularity legitimately enough by use of the reference experiment, as we shall see in the discussion of reinforcement. The circularity arising from defining the stimulus in terms of its effect is perhaps less serious than it seems. That a stimulus reduced to its basic components seems necessarily to involve a gradient or ratio of physical energies and that a precise mathematical description of the stimulus would be a worth-while enterprise is not challenged. Any activity on the part of sensory psychologists in the direction of explicitly and rigorously defining variables would be an important step forward. The only objection is to the notion that generating such analytic geometries would obviate the circularity in definition. Someone at some point, *must* react to the stimulus before it can be labeled a stimulus.

Measurement

The task of experimental psychology is not simply the identification of stimuli and the specification of responses. It is the determining of relationships between variables in the physical environment and the behavior of the organism. Following the practice of mathematicians, we refer to the behavior of an organism as the dependent variable in a functional relationship. This dependent variable, the behavior we are studying, is a function of the independent variable. Before we can develop mean-

ingful statements relating dependent variables to independent variables, we must find a meaningful measure of the two variables in question. Since measurement is almost always implicit in the definition of the independent variable, the quantitative definition of the stimulus seldom presents a problem. Arriving at a meaningful quantification of behavior, however, is something else again. Several ways of measuring behavior have been used experimentally. In the early investigations of the salivatory reflex in dogs, magnitude of response was studied (Pavlov, 1927). Magnitude of response would seem to have all the desirable characteristics to be found in a response measure. One can readily devise instruments to measure differences in physical intensity of activity, whether they record the force with which a rat presses a lever, the sound intensity of a vocalization, or measure the volume of saliva secreted over an interval of time. Under many circumstances, response magnitude is appropriate as a measure of response.

A second measure of response strength is latency, which is the interval of time elapsing between the presentation of a stimulus and the occurrence of the response appropriate to that stimulus. Latency is a popular measure of response strength in studies employing mazes or runways.

A third measure of response strength is the frequency or rate at which responses occur in time. Many arguments, both theoretical and empirical, have been advanced that the measure of response frequency is the most suitable for an analysis of behavior. We shall return to these arguments.

Before passing to a critical analysis of the properties of an adequate response measure, let us consider a

fourth measure: probability of responding. As the concept of response probability is employed by psychologists, it may refer to any one of three different concepts or operations. The first is a naive, unspecified notion of likelihood of action. In this sense it is synonymous with the undefined notion of response strength. A second usage refers explicitly to a mathematical definition of probability of responding, although probabilists are not unanimous in their agreement upon an appropriate definition of probability. The third use of probability is with a relative frequency definition. An example of this is the concept of the *behavior ratio* as employed by Tolman (1932). In statistical learning theory, response probability is defined as the proportion of, say, right-hand responses over a series of trials (Estes and Straughan, 1954). Just as in Tolman's concept, response probability is a behavior ratio. This is an acceptable definition as long as the term is subsequently used only in strict adherence to this definition. It is hard to resist the temptation to use the term to refer to actions that one is about to perform or that organisms have a "tendency" to do. Particularly from the clinician's standpoint, it would be convenient if some measure of response strength or probability were available from which he could predict a unique event. It is probably not reasonable that we should attempt to make predictions concerning such unique events. Rather than speak of the probability of a given event, it would be more useful to state probability limits or bounds for a given set of conditions. Although the prediction of unique events may not be logically impossible, given new ordering assumptions and techniques of analysis, it is not immediately feasible.

The final measuring technique is resistance to ex-

tinction. This is a special case of frequency of responding in which we count the amount of behavior emitted by an organism when conditions appropirate to the maintenance of that behavior no longer exist. In conditioning an organism, the experimenter performs operations that not only will alter the form of behavior under consideration but will also maintain the strength of the behavior. When those operations are discontinued, behavior decreases in strength. The rapidity with which the behavior is extinguished may be taken as a measure of response strength.

As stated in Chapter One, the appropriate method of analysis depends upon the characteristics of the subject matter. Certain criteria, however, do exist to evaluate the adequacy of a given measurement of behavior. There should be a reasonable range of variation; it should be possible to measure changes in behavior over a wide range of values. The property of response to be measured should be sensitive to the class of variables being studied. Although it is a reasonable problem to determine whether a given response measure is related to a particular variable, one does not generally select a response measure that is unlikely to vary as a function of experimental manipulations. There is little point in measuring an aspect of behavior that is unchanging irrespective of any experimental manipulations which may be performed upon the organism. On the other hand, the particular characteristic of behavior we measure should not be so labile that it flips around as a function of any and all variables. If the measure is extremely sensitive to many variables, the experimenter is challenged to contrive a setting where irrelevant or extraneous influences are held to a minimum.

A problem to which a great deal of attention has been given in the literature on the philosophy of science is the distortion of a subject matter by the act of measurement or observation. The uncertainty principle and the law of complementarity have been extended into arguments of behavioral analysis (Skinner, 1950). The first two measures of response strength, magnitude and latency, are easily distorted by measuring procedures. In particular, magnitude is distorted by conditioning. Skinner argues that rate of responding, response frequency, is the technique of choice, and that rate of responding is operationally equivalent to probability of response. Estes (1950) has demonstrated mathematically that given certain limiting assumptions, rate of response is directly proportional to mathematical probability. For the layman, response frequency has a happy overlap with the vaguely defined concept of response strength.

Two comments about response magnitude must be made. Magnitude may well be an appropriate measure of response strength in many cases. It may not be more distorted by the process of measurement than any other property of responding that an experimenter might select. It is true, especially in organisms that have had extensive social contact with others, that response magnitude is quickly subjected to differential reinforcement that alters its range of variation. For example; we observe that when an infant is hungry it cries more loudly than when it is not. It might, in fact, be found that in general the longer the period of food deprivation, the louder the child cries. Even in an older child who has a vocabulary appropriate to the acquisition of food one often finds that the vocal response becomes louder with

increasing deprivation. However, parents characteristically train a child to request food and other objects in a well-modulated tone of voice, prefaced by certain socially acceptable conventions, such as saying "please." Once the intensity of behavior has been modified in this way, that particular aspect of responding has become less appropriate as the subject of measurement of response strength. Not only does it not really represent whatever we naively refer to when we speak of response strength, but it no longer has a wide range of variability.

The second comment is that response magnitude is a summary measurement made over a period of time. It can be argued, for example, that the amount of saliva that Pavlov's dogs secreted to a tone might just as well have been treated as a rate or frequency measurement, since the amount of saliva collected at the end of an interval of time was the sum of individual salivations occurring during that time. In order for a larger quantity of saliva to have been secreted at the end of the interval, the frequency at which drops of saliva were secreted had to be greater. The instrumentation required to demonstrate this argument with regard to skeletal muscular activity would need only to show that the more intense or the more forceful action is mediated by a greater frequency of neural and muscular activity. Magnitude is not conventionally analyzed in this way for two reasons. It is simpler to measure force or volume, and measurement of rate changes over such short intervals of time would be not only an uneconomical but possibly an uninteresting type of data to gather. Although under certain circumstances response magnitude may be a perfectly acceptable laboratory measurement, it is known

that it does not vary in a 1:1 relationship with other measures. One measure can not always be substituted for another.

Arguments against latency as a response measure also run along the line that the measure is not general to the on-going behavior of the organism. It is true that the environment does not operate to permit observation of the precise instant a controlling stimulus occurs. The interval elapsing between the occurrence of that stimulus and the occurrence of the appropriate response usually cannot be measured. Generally, in dealing with operant behavior, the crucial stimulus is unspecified or even, perhaps, unspecifiable. It is the strength of the position that it is possible to establish valid functional relationships in the absence of such stimuli. A second objection to latency rises from the fact that latency seldom is a measure of single response classes. It may be contaminated by preparatory activities in which the organism engaged. Other questions may be raised concerning latencies observed in mazes and runways as they have traditionally been employed. For example, is the position of the organism in a runway as a function of time of primary importance or is, perhaps, the rate of change of position of greater importance? In experiments using runways where latency of response has been taken as the measure, it has been observed that the animal runs faster the nearer he approaches the goal box at the end of the runway. The change in the speed of the animal as he traverses the runway has been used as a measure to define a concept known as the goal gradient. The change in rate of running has occupied a position of central importance in learning theory. It can equally well be argued that the goal gradient is an artifact of this particular re-

sponse measure. While it is true that the animal moves more rapidly toward the end of the runway, the accelerative force may remain constant or even decrease. Although we do not intend to undertake a critical evaluation of the goal gradient hypothesis at this point, the example is presented to illustrate that a given measure may mask subtle properties of the situation which may be important.

That latency is not a good measure of a single response class is further accentuated when a maze is used. Although one may deliberately elect to investigate response classes of varying degrees of heterogeneity, one should not unintentionally contaminate a response class. One can argue that a right turn is a right turn no matter how it is accomplished, but a maze maximizes the likelihood that irrelevant and extraneous influences will distort or disrupt the behavior under observation.

Arguments favoring the rate measure against alternative measures have considerable relevance, but the rejection of those alternatives does not insure that there are no arguments against rate. Some of the objections to magnitude and latency of response may also be raised against frequency of response. Rate, too, is subject to modification by controlling variables. This, in fact, is the basis for the entire area of research on schedules of reinforcement. The analysis relating the rate of response to a mathematical concept of probability deserves considerably more careful scrutiny than has heretofore been given it.

CHAPTER THREE

Basic Conditioning Processes

It is the very nature of learning that the behavior of an individual must be changed. Whether the change involves the acquisition of new response modes or the strengthening of behaviors preexisting in the individual's repertory, some behavior must be strengthened. We refer to the process of strengthening as *conditioning* and to its means of accomplishment as *reinforcement*.

The term *reinforcement* means strengthening. An army unit is reinforced when it is strengthened by the addition of more troops. Concrete is reinforced by the addition of iron rods. Behavior is reinforced by appropriate strengthening contingencies. It is tempting to equate reinforcement with "reward," "pleasure," and other hedonistic concepts. Quite often operations that reinforce behavior fit such labels, but identification of the reinforcing process with reward is dangerous. It attributes a surplus meaning to reinforcement which implies processes for which there is no evidence and which are based upon the prescientific conception of man.

A reinforcer, then, is just that which reinforces. This

is a circular statement. Circularity, however, is obviated empirically by the reference experiment. From a priori considerations one does not know that a given operation will reinforce behavior. One performs the experiment and observes that an increase in response strength does occur as a result of the operation. The operation has reinforced the behavior. Having labeled the operation as reinforcing, one then predicts in a different experiment involving a different response class that this operation will also increase the probability of that response class. One can then verify that the particular operation does in fact act in this way; the operation, as a reinforcer, has trans-situational validity. Other effects of reinforcement emerge, depending upon the way in which reinforcement is made contingent upon behavior. A response that has been reinforced decreases in strength when the reinforcer is withheld. In addition, there are characteristic effects upon topology of behavior, frequency of behavior, and so on, as a consequence of specific contingencies between reinforcement and response.

It is worth mentioning in connection with our discussion of reinforcement that, historically, concern was voiced over the way in which an event could act backwards in time. It was of particular concern in what has been called instrumental or operant conditioning, where the reinforcing operation follows the occurrence of a response. This particular problem need no longer concern us; it was only a problem in the first place because of faulty analysis. Reinforcement is the result of a physical event acting upon the organism. Food is not given to the response, but to the organism; the effect upon the organism is an alteration in its behavioral topology. There is no question of action backwards in time.

Respondent and Operant Conditioning

Following Skinner's terminology, we make a distinction between two general classes of behavior, *respondent* and *operant*. Whether these constitute essentially different types of behavior, involving different processes based upon differing neural structures and effectors, need not concern us here. The two classes of behavior do differ operationally in the manipulations used to control them. Reinforcement is used with both classes of behavior.

Reinforcement of respondent behavior consists of the presentation of two stimuli closely associated in time. In this case, the problem of a surplus meaning of reinforcement does not usually arise because the operations involved do not appear to be "rewarding." The facts of Pavlovian conditioning, or respondent conditioning, are well known, but it is worth while to recapitulate for a moment in order that we may later draw some parallels between these processes and others. A simplified Pavlovian paradigm is presented in Figure 1, showing the pairing or association in time of the so-called unconditioned stimulus with a neutral stimulus. The unconditioned stimulus is one that has a high probability of eliciting a response. The neutral stimulus, in the absence of specific training history, does not normally elicit the response in question. After a number of trials where the two stimuli are presented in close temporal contiguity, the neutral stimulus evokes the response. It can now be called a conditioned stimulus. This is respondent conditioning. If only the conditioned stimulus is thereafter presented repeatedly, the strength of the response evoked by that stimulus decreases as a function of trials. This

decrement in response strength is called *extinction*. When the unconditioned stimulus is again associated with the conditioned stimulus, response strength to the conditioned stimulus is again raised, and it is said that the response has been reinforced.

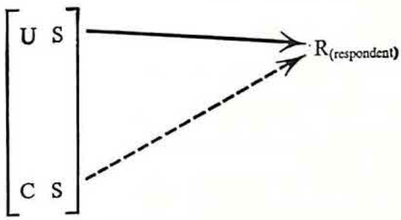

FIGURE 1. Paradigm of respondent conditioning. The conditioned stimulus (CS), for example, a tone, comes to evoke the response, salivation, which is normally elicited by the unconditioned stimulus (US), food, through the repeated temporal association of CS and US.

In operant conditioning the procedures are somewhat different. One does not present an eliciting stimulus to the organism. Instead, one observes the on-going behavior of the organism, and when a member of the selected response class appears, the reinforcing operation is performed immediately thereafter. For example, as the rat moves about the cage and presses a lever, a food pellet drops into the magazine beneath the lever. The rat eats the food pellet, and the observed effect is an increased frequency of pressing the lever. As in respondent behavior, should one allow the organism to continue making the response and not reinforce the response, the strength of the response decreases. It is in operant behavior that one is tempted to label reinforcers "rewards"

and implicitly, if not explicitly, invoke hedonistic explanations to account for learning.

It should be borne in mind that the distinction between operant and respondent behavior is always operational. The distinction rests upon the controlling procedures and not upon other criteria. It has been noted that respondent behavior is frequently mediated by the autonomic nervous system and involves smooth muscles and glands as effectors, and that operant behavior usually involves the central nervous system and skeletal musculature. Such a neat categorization based on anatomical structure would be convenient, but it is not accurate. The same structures may be controlled by both operant and respondent procedures. Leg flexion in the dog, for example, may be controlled by respondent procedures, but it may also be controlled by operant procedures. In this particular example an additional difficulty exists in demonstrating the existence of the required autonomic innervation. From the layman's point of view, of course, the types of activity normally mediated by respondent procedures have been called "involuntary" actions, whereas those controlled by operant procedures have been called "voluntary." In an analysis which does not permit voluntarism, such a distinction is worse than meaningless. Certain theoretical analyses argue that these activities are actually two forms of the same process and that the underlying process is identical in both cases.

In the interest of reducing the number of categorizations with which we must deal, it is worth while to examine some of the distinctions frequently made in texts dealing with the learning process and to abolish some labels that seem to have little functional merit. First of all, it should be explicitly stated that respondent be-

havior is identical to that which has been labeled Pavlovian or classical. Respondent behavior plays a large role in what is usually called "emotion." Operant behavior subsumes the following: instrumental conditioning, trial and error learning, verbal conditioning, motor learning, problem solving, concept formation and insightful solution to problems. This lumping of various categories into one is not an attempt to oversimplify, but is rather an attempt to group various subject matters together in terms of their controlling operations. It is readily seen in the second group that reinforcing operations are of the operant variety. In elaborating the concept of reinforcement, we will find that it is possible to simplify the situation further. The various labels that abound in the field of learning have arisen historically because many individuals have converged upon the subject matter of behavior from many different starting points and have brought with them concepts originating in these various starting points. Little attempt has been made to coordinate procedures or to reduce the number of concepts within the field. Although no one would deliberately have encouraged the proliferation of categories, their number has grown as a natural consequence of the exposition of differing points of view.

Reinforcers: Primary, Secondary, and Generalized

Given certain deprivation states of an organism, there is a class of stimuli that will, without any prior history of training, reinforce operant behavior. These are the *primary reinforcers*. They are not primary because of any

biological primacy, but because they do not require a prior training history to be effective. It may be that they also have a biological importance to the organism above and beyond other classes of stimuli, but to impose this as a condition of definition implies theoretical assumptions we are unwilling to make. The distinction between primary and other reinforcers of operant behavior does not depend upon such theoretical assumptions.

$$R_{(\text{operant})} \longrightarrow \begin{bmatrix} S^R \\ \\ S^r \end{bmatrix}$$

FIGURE 2. Paradigm for the establishment of a secondary reinforcer (S^r). The neutral stimulus, say, light, acquires reinforcing properties through repeated temporal association with a primary reinforcer (S^R), food.

As in Pavlovian conditioning, it is possible to alter the function of stimuli in such manner that those which are normally not reinforcers will become reinforcers through association with other stimuli. If one associates, say, a click or a flash of light with the presentation of food in operant conditioning, one observes after a number of pairings of light and food that the light itself will become reinforcing. The light becomes a *secondary reinforcer.* The paradigm for establishing a secondary reinforcer is shown in Figure 2. It has been noted that if the effectiveness of a secondary reinforcer has been established under conditions of water deprivation, that secondary reinforcer will be most effective in the future when the ani-

mal is water deprived. If it has been established under conditions of food deprivation, it will be most effective for an animal that has been deprived of food. This somewhat limits the possibilities of control by secondary reinforcers. Fortunately, however, the range of control can be extended.

If the neutral stimulus is associated with several primary reinforcers under a wide range of deprivation conditions at different times, then that stimulus will come to be effective under all of those conditions. Such a broadly based secondary reinforcer is called a *generalized reinforcer*. It does not depend upon a specific deprivation condition for its effectiveness. Experimentally it is probably not possible to establish a pure secondary reinforcer because it is not possible to establish a single state of deprivation in an organism. An animal that is food deprived does not consume as much water as one that is not food deprived. Similarly, a water-deprived animal does not ingest as much food during water deprivation as does an organism with unlimited access to water. The result is a confounding of the deprivations to the extent that any secondary reinforcer is necessarily established in the presence of more than one deprivation state (Verplanck and Hayes, 1953). A number of theoretical discussions would have met with the ignominy of nonpublication had this fact been learned somewhat earlier. To the extent that the experimenter deliberately manipulates deprivation conditions and associates a neutral stimulus with a wide range of primary reinforcers is the reinforcer more generally applicable to a wide range of conditions.

Certain similarities between this process and the process of respondent conditioning are evident. Figure 3

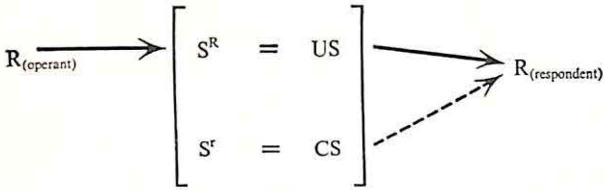

FIGURE 3. Schematic representation of the relation between the establishment of a secondary reinforcer of operant behavior and a conditioned stimulus for respondent behavior.

combines the two processes. Stimuli may have a dual function in that they may be effective in affecting the strength of both respondent and operant behavior simultaneously. It is possible that establishment of a secondary reinforcer of operant behavior depends upon a stimulus having acquired control of respondent behavior. An experimenter may focus primarily upon operant conditioning, but he is not dealing with an organism that is in a respondent vacuum. Similarly, though one may be primarily concerned with respondent reflexes, and though the organism may be standing quietly in a harness, the on-going activity of the organism necessarily includes a very wide range of operant behavior.

Reinforcers: Negative and Positive

Another set of distinctions among reinforcers of operant behavior cuts across the preceding definitions. In one case the *presentation* of a stimulus will strengthen behavior; an example is giving food to a food-deprived rat for pressing a lever. The effect is to increase the strength of the lever-pressing response. On the other hand, the

removal of a stimulus can strengthen behavior. For example, by pressing a lever a rat placed on an electrified grid can terminate the electric shock for an interval of time. This also results in an increase in lever pressing. We refer to these two classes of reinforcers as *positive* and *negative* reinforcers, respectively. Again, to avoid problems of surplus meaning, "positive" and "negative" are defined explicitly to refer to additive or subtractive laboratory procedures. In positive reinforcement the stimulus is *added* to the stimulus matrix. In negative reinforcement the stimulus is *subtracted* from the stimulus matrix. This distinction unfortunately is not universally clear in experimental literature.

In the preceding examples of secondary and generalized reinforcers, primary reinforcers have been of a positive type. This need not always be the case. It is possible to associate a tone with an electric shock so that *termination* of the tone alone will strengthen a response. One can again reduce the number of categories in the literature on experimental conditioning. Under the heading of negative reinforcing procedures one must include procedures labeled "escape" conditioning and "avoidance" conditioning. These experimental procedures involve nothing more than ordinary operant reinforcing operations in which negative reinforcers are employed. Avoidance conditioning involves the use of a "primary" negative reinforcer; escape conditioning entails the establishment and use of a "secondary" negative reinforcer.

In considering negative reinforcement one must distinguish between negative reinforcement and *punishment*. Frequently these terms are misused synonymously. No good term suggests itself to categorize the class of

objects that can serve as positive reinforcers. In the case of negative reinforcement, however, the term *aversive* is used to describe a class of stimuli that can serve as negative reinforcers. By definition, an aversive stimulus is one which, other things being equal, will cause an organism to act in a manner that removes him from that stimulus. An aversive state of affairs is employed *both* in negative reinforcement *and* in punishment, but the response-stimulus contingencies are different. Punishment and its effects will not be discussed in detail, but it is appropriate to introduce some important distinctions at this point. The distinction between reinforcement and punishment rests upon the temporal contingencies between response and stimulus. If an aversive stimulus is present, the organism responds and the stimulus is immediately removed; the effect is a strengthening of the response. This is negative reinforcement. However, if the stimulus is not present, the organism responds and the stimulus is immediately presented; the effect is punishing. It is unfortunate that there is no other term for this latter procedure and its effects, since the term *punishment* carries with it surplus meanings, as does *reward*. It would be better to call the aversive stimulus so used a *suppressor*. Figure 4 is a fourfold contingency table with two categories of stimulus on the abscissa. The ordinate indicates that reinforcement occurs from the addition of stimuli of Type A and by the subtraction of stimuli of Type B. Symmetrically, it is seen that suppression or punishment occurs as a consequence of presenting stimuli of Type B and by the removal of stimuli of Type A. The latter is exemplified by taking candy from a baby.

Considerable argument has taken place as to whether the concept of operant reinforcement is ade-

	Stimulus Class	
	A	B
Reinforcement	+	−
Punishment	−	+

FIGURE 4. Relation between positive and negative reinforcement and punishment in terms of the class of stimulus employed. + indicates presentation, − indicates withdrawal of the stimulus.

quate to explain the learning of complex human behavior. In particular, some studies (Wallach and Henle, 1941) indicated that some material reported by Thorndike was limited in that operations purported to be reinforcing did not, in fact, always reinforce. For an experimenter to say the word "right" to a subject does not always reinforce the subject's behavior. Two remarks are in order. One, the situation is not as serious as Henle believed it to be; two, it is far more extensive than the critics believed it to be. Giving a rat a food pellet for pressing a lever generally reinforces the rat's lever pressing. However, this is true only under the appropriate deprivation conditions. For primary reinforcers to be effective, it is necessary to engage in some initiating operation so that the reinforcer, in fact, reinforces. As stated before, every situation is special. The situation where primary reinforcers reinforce is special in that it depends upon the establishment of certain prior conditions. There is in the study just cited

no contradiction of the facts of reinforcement, but rather a confirmation of the general principle just stated. To state a functional relationship, one must consider other relevant status variables. The fact that these status variables exist does not vitiate the relationship whatsoever. It is really only surprising that this state of affairs has been cause for surprise.

The Drive Reduction and Contiguity Theories

In conclusion, it is appropriate to examine briefly two major theoretical positions. Detailed expositions of these points of view may be found both in primary works and in excellent secondary critiques. However, it is relevant to sketch the conditions these positions argue to be necessary for the conditioning process. First, let us consider the drive reduction position (Hull, 1943, 1951, 1952) which argues that conditioning depends upon a "need state" existing in the organism. "Need state" is an inferred condition that supposedly operates upon a secondary inferred event or state called a "drive." Operations that reduce drive are, per se, reinforcing. Drive reduction must occur at some stage so that learning may take place.

The second theoretical position is that of contiguity theory (Guthrie, 1935), which states that reinforcing events serve only to protect the last response made in a temporal series by altering the stimulus circumstances so that the stimuli being conditioned are removed or altered. Removal prevents competing behaviors being conditioned to these stimuli. The considerable literature re-

porting attempts to decide between these theoretical positions has perhaps been foredoomed to failure. Consider a Venn diagram as in Figure 5, where the class of events that constitute changes in environment in contiguity theory subsume classes of events that represent drive reductions in the alternative system. It is clear that it is impossible to reduce a drive without altering the stimulus situation. It is logically possible to determine whether or not there are changes in the environment that are not drive reducing. However, the number of

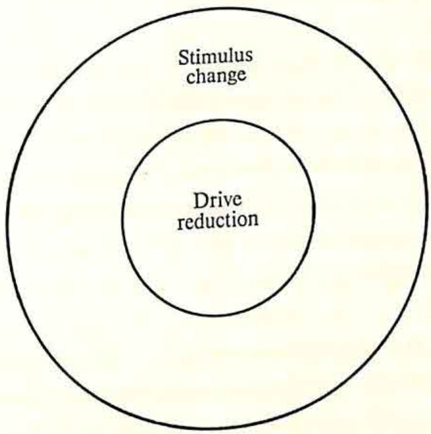

FIGURE 5. Representation of the relation of stimulus change associated with drive reduction to more general stimulus change.

hypothetical states or variables allowable in the various drive reduction systems very nearly eliminates all possibility of a conclusive test between opposing theories. The two positions predict essentially the same things; at the level of simple conditioning, there is no basis for preferring one to the other. Inadequacies emerge as one moves further into more complex processes.

When people begin to study a particular phenomenon, they carry into their study the language and concepts that they had previously learned were appropriate to that subject. This is what happened in the early investigations of learning. Just as there is a label, "learning," to apply to the process of behavioral acquisition, there is the polar term, "forgetting," which presumably refers to an opposite process of behavioral loss or decay. For the layman, forgetting is the reverse of learning. Unfortunately, the experimental situation has shown that the matter is not quite so simple. Furthermore, in the English language the antonym of "reward" is "punishment." Punishment was therefore thought to have an effect opposite to that of reward. We recognize reward as a poorly defined lay term referring to those operations technically called reinforcing. We also recognize procedures that we call punishment. However, the effect of punishment is not simply the opposite of the effects of reward or reinforcement. The effects of punishment are quite complex.

Thorndike, in the original law of effect (Thorndike, 1913), spoke of a noxious event, punishment, as exerting an effect opposite to that of reward. A reward, according to Thorndike, strengthened bonds; punishment weakened those bonds. With the publication of the truncated law of effect (Thorndike, 1932), he recognized that punishment did not exert an effect opposite to that of reward. His view that punishment was a symmetrically reverse operation to reinforcement was revised. The effects of punishment have been explored in detail, and are now well understood (Estes, 1944). The question remains, however, whether there is a process which does have an effect opposite to that of reinforcement. A possible an-

swer lies in extinction, where behavioral strength is seen to decrease with repeated elicitations or evocations of the unreinforced response.

Extinction

It is worth while to examine some of the theoretical analyses of extinction to gain some insight into whether extinction is a reverse operation to reinforcement. The interpretation of extinction proposed by drive reduction theorists is that extinction results jointly from the accumulation of fatigue products produced by work and from stimuli which—through their association with fatigue products—have become conditioned to effect behavior in the same way fatigue products do. According to this view, behavior persists only where the accumulation of work decrement, fatigue product, or inhibitory potential is overbalanced by the effect of reinforcement, the accumulation of habit strength. Serious objection is raised to this formulation by the effects of schedules of reinforcement, because the preceding argument should predict that where more work is required of the organism per reinforcement, behavior should be in least strength. Behavior should rapidly extinguish where only occasional responses are reinforced in contrast to the situation where each response leads to reinforcement. In the latter case, the balance between the effects of reinforcement and the effects of work decrement is more favorable than where more behavior is required per reinforcement. The data are unequivocal in demonstrating that the contrary is the case. In fact, response strength is

greater as the ratio of behavior required per reinforcement is increased.

Alternative explanations have been advanced in defense of the drive reduction position in the form of the response unit hypothesis. This hypothesis argues that what is reinforced under a schedule is not, for example, an individual bar-press response but rather patternings or groupings of bar presses, and that the larger unit then undergoes reinforcement. Consequently, if one counts the larger unit, the effect of reinforcement is the same or perhaps not quite so favorable as in the case of continuous reinforcement. However, this is begging the question because the effects of reinforcement must in some way be enhanced by requiring more work; it is an inescapable fact that more work is required under a schedule. The actual output of energy is greater under schedules than it is under continuous reinforcement. In order that even equal strengths of behavior be maintained by the two conditions, the output must in some strange way have an effect of such a nature that reinforcement is enhanced by greater effort. This bends the theory to fit the facts in a way that is less than ideal from the standpoint of theory construction.

The second theoretical interpretation of the extinction process is taken from contiguity theory. To understand extinction according to this position, one must refer to the treatment of the conditioning process. When a stimulus and a response occur together in time, conditioning is accomplished; when the stimulus is presented again, the response that was associated with it previously will be called forth. Conditioning is complete on the first pairing of the response and the stimulus. A casual reading of these statements might lead one to assume that learn-

ing is instantaneous. This would be disturbing since it is obvious even to the most naive that learning is gradual and often arduous. However, contiguity theorists mean something quite different when they speak of response and stimulus than is implied by a gross definition. The effective stimulus is conceived of as being subdivided into elements. A response is said to be conditioned to a certain proportion of the stimulus elements available to the organism: that proportion which the organism samples, perceives, discriminates. By successive exposures and sampling, the organism gradually exhausts the total set of available stimulus elements.

The probability of response approaches 1.00 as an asymptote, because response probability is defined to be the proportion of conditioned elements to the total available set. Response probability increases gradually as the supply of available unconditioned elements is exhausted, that is, as elements are transformed from the unconditioned to the conditioned state. Reinforcement plays no part in the learning process except as an artifact. If an organism is placed indefinitely in an environment where he performs many different actions, and if each action is therefore immediately conditioned to those elements present, it is seen that a great deal of unconditioning or deconditioning must occur. Response (n) ties up a certain number of stimulus elements that are conditioned to it, but response $(n+1)$ in turn becomes associated with some of those same stimulus elements. An element, by definition, cannot be conditioned to two incompatible response classes; an element either is or is not conditioned to a particular response. If it has been conditioned to a response (n) and then is conditioned in turn to response $(n+1)$, it can no longer be conditioned to response

(n). In other words, the conditioning of the element to response ($n + 1$) has interfered with its association with response (n). Those operations which reinforce behavior are regarded as changes in the situation and, according to contiguity theory, they are essential for learning to take place. Change in the stimulus situation alters that situation so as to protect the last response. In other words, if the last response were response (n), then response ($n + 1$) could not take place in the presence of the same stimuli. The situation has been so altered that if response ($n + 1$) occurs, it occurs in a *different* stimulus setting and does not interfere with conditioning that has gone on in the old setting. The last response, the response that alters the environment, is the response that is least interfered with and which gradually becomes the only response made in that setting. Antecedent responses in the setting undergo progressive interference from later response instances. The earlier actions the organism performs, which are unsuccessful in altering the situation, gradually suffer interference by subsequent actions. In comparing contiguity theory with drive reduction theory, one may regard as part of the stimulus situation stimuli arising from stomach contractions in a rat that has been deprived of food. The rat smells in the corner of the box, scratches himself, climbs upon the wall, peers about; yet all of these actions do not alter the stimuli arising from stomach contractions. Eventually he presses a lever which delivers a pellet. If the rat ingests the pellet, then the last response before ingestion will be least interfered with since the ingestion of the pellet has altered the situation sufficiently to protect the response. Earlier responses all suffer interference by succeeding responses, up to the last

one. As the rat presses the lever on succeeding occasions, preceding elements of the response chain drop out through interference; the intervals between lever presses become shorter, and, inversely, the rate of lever pressing increases.

In the contiguity treatment of the learning process, no assumptions are made concerning the pleasantness of the reinforcing event for the rat, the independent properties of the reward, or the necessity for drive reduction. It is, however, a theoretical interpretation of the learning process. It is not a strictly descriptive analysis, and although the argument is plausible enough, important problems remain. For example, how much of a change is required in the stimulus field for the change to protect a response? In other words, how large a change is a change? The environment is constantly changing in time, constantly in flux. Yet some event must occur, some event discriminable to the organism and produced by that organism's action, which protects the response that produced it. What is the nature of such a change? This is only one of several basic problems remaining for contiguity theorists to resolve. To return to the contiguity treatment of extinction, extinction is seen to be no more than interference in behavior by other succeeding behaviors. The contiguity theorist regards conditioning symmetrically with extinction in that conditioning is but a transfer of response probabilities from one class to another. If one categorizes the behavior of an organism into class A and $\tilde{A}$, then conditioning of response A is the extinction of response $\tilde{A}$, and vice versa. At the level of theoretical analysis, extinction may be viewed as the reverse of the coin of conditioning. The conditioning of

one class of competing behavior is the extinction of another, and the conditioning of that second is the extinction of the first. Conditioning and extinction are aspects of the same process, namely, the effects of interference or protection of a given class of behavior.

As stated before, it is impossible to choose between the Hullian and the Guthrian theories of learning on the basis of what each predicts in simple conditioning. Those things which are drive reducing in the Hullian context inevitably are also changes in the environment in the Guthrian. For a more satisfying answer to our question as to whether there is an operation inverse to the reinforcement operation, we are, as always, to fall back upon experimental evidence. Let us take as our initial measure of the steady state of an organism's behavior the operant level of that behavior, the level at which, for example, a rat presses a lever. This is the rate at which he responds before ever having been reinforced for pressing the lever. After conditioning the lever pressing until the rate of response is high, we may then extinguish the response until the rate is quite low, lower, in fact, than it was in the initial steady state, lower than the operant level. Have we then succeeded in reversing the conditioning process? Have we succeeded in lowering the initial probability to the level at which it existed before the animal's response had ever been reinforced? If one looks only at the rate of response, one might answer yes to this question. However, if we undertake to recondition the particular response in question following extinction, we discover that it is quite easy to recondition that response. It is easier to recondition it following extensive extinction than it was originally to condition the response. This ease of reconditioning as contrasted with original condition-

ing is evidence that the extinction did not, in fact, return the organism to the original state. Extinction did not clear the organism. On this evidence, the extinction process does not look as though it was the exact reverse of the original conditioning. Extinction does not seem to succeed empirically in reversing the course of original learning. It has been stated in comparing the human organism to a computing machine that the human organism can only be cleared at death (Wiener, 1948). The evidence of the effects of extinction as a clearing procedure would seem to support this notion.

It has been argued by some that extinction procedures of the type just described have not gone far enough. For example, in the contiguity analysis it is stated that the environment is continually changing and that not only must the organism exhaust the immediately available set of elements through counterconditioning or extinction, but it must also exhaust the entire set of previously conditioned elements. Time must elapse so that elements not immediately available but which have been conditioned at an earlier time may reappear (Estes, 1955). Such elements have passed out of the immediately available stimulus set through physical changes. They are conceived of as diffusing in and out of the available set at a given rate with fixed probabilities. In order to extinguish all elements, extinction must be pursued until all elements not presently available to the organism have an opportunity to reappear; then extinction takes place to these. The probabilities of recapturing, resampling every single stimulus element to which an organism may have been conditioned are remote. Thus much must be granted the contiguity theorist. This, then, is a quite plausible accounting for the facts.

Forgetting

The processes of extinction are of particular relevance to education, but we must clarify confusions that often arise in treating complex human behavior in terms of the present analysis. First, extinction should not be equated with forgetting. Forgetting is a real phenomenon that is all too frequently observed in students, but it is a more complex combination of processes than simple extinction. At least two, and perhaps three, factors may be involved in forgetting. The first of these may be extinction. A behavioral class may be conditioned to a set of stimuli, but the conditioning then undergoes interference through the subsequent conditioning of other behaviors to stimuli that compete with the original conditioning. This process has already been discussed in treating the contiguity analysis of conditioning. If behavior is conditioned to a set of stimuli but is not protected from subsequent conditionings, the subsequent conditioning will interfere with the earlier. To the extent that this analysis of extinction is valid, it applies to some instances of forgetting.

A second condition producing forgetting is where behavior has been adequately conditioned to a high probability of occurrence under the original conditions of learning but where the necessary stimulus support is lacking at the time behavior is subsequently required. This condition is similar to that in the first example but does not call for specific counterconditioning; it may simply be that necessary components of the environment that were present during learning are not available at the time performance is demanded. Students have this experience when they have prepared for an examination

in a specific place, such as their dormitory rooms, but are unable to remember information at the time of examination. When they return to the dormitory room, they are suddenly able to recall the forgotten answer or the forgotten material.

A third variable that contributes to forgetting has to do with the effects of aversive control. Forgetting produced by such conditions is called motivated forgetting. As pointed out earlier, the operant behavior of an organism does not occur in a respondent vacuum. If the situation in which behavior occurs has aversive properties causing the respondent behavior to become prepotent, the organism may simply be so heavily engaged in respondent behavior elicited by the aversive properties of the environment that he has no time to emit the particular operant behavior in question. Often the operant behavior itself leads to aversive consequences of a type that condition respondent behavior. We must not lose sight of the fact that a response may be a stimulus not only to other organisms but also to the organism that emits the particular response. Stimulus feedback resulting from an operant response can acquire conditioned aversive properties through association with aversive consequences that are mediated by the physical or social environment. When the individual begins to emit the chain of behavior that leads to aversive and conditioned aversive effects, these effects may suppress that behavior itself before the chain of responses is complete. In this way actions may become anxiety producing to the extent that the individual is incapable of completing the action. Being unable to complete the action may involve being unable to realize some overt accomplishment, being unable to talk about the action or, what would be of still more

relevance to the psychotherapist, the individual may even be unable to recognize that he has begun a chain of behavior that has aversive suppressing effects. In this latter case, the individual may be so unaware of his own behavioral processes that forgetting is not the proper term to apply. In such a case, forgetting is of such magnitude that it infringes upon psychopathological phenomena.

We observe in our own experience that forgetting occasionally seems to be a passive thing. Certainly, the type of forgetting exemplified in the second case just discussed, where forgetting takes place because the necessary supporting stimuli are ineffective by their absence, can legitimately be regarded as a passive process, if, indeed, as a process at all. *Extinction*, however, is *not* a passive process. In the literature on extinction one does find discussions arguing for the view that extinction constitutes a breakdown of associations over time in much the same fashion that radioactive materials decay. Experimental evidence, however, forces us to reject this interpretation out of hand. Extinction does not occur in the absence of behavior, be it respondent behavior or operant behavior. If a response has been conditioned, and if for some reason the organism has been unable to make that response, then when the organism is again placed in the situation where he can make the response, no loss will have occurred in the behavior. The response is still of great strength. On the other hand, if the organism is placed in a situation where the response can be made but where, if it occurs, it goes unreinforced, then extinction occurs. Extinction is an active process.

In view of this, it is clear that interference with the emission of behavior as produced by the introduction of

aversive techniques of control only delays the course of extinction. In order for extinction to occur, the response must occur. If any manipulation is employed which prevents the response, extinction will be deferred accordingly.

CHAPTER FOUR

Motivation

Perhaps a more accurate title for this chapter would be "Problems of Motivation," for it is not clear that the area of study commonly referred to as motivation is concerned with a set of behavioral processes apart from those which we have already discussed. The problem of student motivation is, however, of primary importance in the technology of education, irrespective of the nature of the underlying process or processes to which the term refers. All living organisms behave all the time. Even sleep is behavior. One problem of motivation is to establish more active classes of behavior. We cannot modify behavior unless behavior occurs. A second problem of motivation with which we must be concerned is to insure that reinforcers be effective, first of all, and, secondly, that they be relevant to the class of behavior we wish to develop in the learner.

In the discussion of reinforcement in Chapter Three it was pointed out that for reinforcement to occur, even primary reinforcement and even though the reinforcer has biological relevance, it is necessary for there to have

been some antecedent operation to which reinforcement is complementary. Food is reinforcing to a hungry dog—not to a satiated dog. One function of a motivating procedure is to insure that there be effective reinforcers available for subsequent use.

The other function of motivating procedures and one with which we are more familiar has to do with stirring the organism to action. There are two general procedures by which such stirring to action is accomplished in the experimental laboratory. Let us turn to these for review and for a consideration of what meaning these procedures have for the practical control of human learning.

Deprivation and Aversive Stimulation

The first procedure is deprivation. An organism such as a rat may be deprived of food for a length of time, say 24 hours. We observe that at the end of this period of deprivation the rat is very active. This is a primitive response, an innate response of this particular organism to this particular condition. The use of a genetically determined pattern of behavior increases the likelihood that the organism will give us some approximation of the kind of behavior we intend ultimately to establish in his repertory. As long as the animal lies quiescently in a corner, contentedly scratching fleas or dozing, reinforcement will only succeed in establishing behaviors that compete with his resting performance. Moreover, if the animal is in such a quiescent state it is probably satiated—and it is then likely that the operation which we hope

will be reinforcing will not prove to be reinforcing at all.

A second technique has to do with the use of aversive stimulation, such as electric shock or other strong stimuli. If we place a rat in a box, the floor of which is an electric grid, and send a mild electric current through the grid, we observe that the activity level of even a satiated rat increases. Perhaps it is this fact that makes the use of aversive motivation attractive to a teacher in a practical situation where he neither has control nor perhaps even knowledge of the particular deprivation states of his students. The use of aversive motivating devices has its disadvantages, as might be anticipated from what we know of the effects of punishment. As we have already seen, one of the effects of aversive stimulation is the elicitation of strong respondent activity. Respondent activity may be so strong that it is prepotent over operant behavior. The teacher who utilizes strong aversive procedures in the classroom may generate in his students emotional predispositions that actively interfere with the emission of the desired operant behaviors. Another byproduct of this form of control with which we are all familiar is that the teacher, himself, being identified with his procedures, becomes a conditioned aversive stimulus. As such, the teacher loses whatever potential he may have possessed for becoming effective—or for his behavior becoming effective—as any kind of positive generalized reinforcer. The student spends so much time attempting to reduce the anxieties or tensions produced by the aversive situation and trying to solve the interpersonal problem in emotional or respondent terms that he is handicapped in facing and dealing effectively with the substantive tasks at hand. Countercontrol on the part of

the student to such behavior on the part of the teacher is the most frequent byproduct of aversive control. Even where countercontrol is impossible, the general consequence is a lessening of the effectiveness of the teaching situation.

Certain theoretical problems must be acknowledged at this point. These problems are seldom explicitly faced in the literature on the subject and, to this writer's knowledge, no solution has yet been found for some of the confusion that exists in this area.

The perceptive reader will already have noted that we have referred in three separate contexts to the use of a technique of withholding what is or what might potentially be reinforcement. In the first instance, the withholding of reinforcement leads to extinction. In the second instance, withholding of reinforcement has been likened to the presentation of an aversive stimulus as punishment. Thirdly, in this chapter we have begun our discussion of motivation by referring to an operation called *deprivation*. In all three cases, objects or events that are potentially reinforcing by their presence are withheld. We have already encountered an analysis which treats the behavioral phenomena of extinction as transfers of probability of response from one class to another. In these terms, can a meaningful distinction be drawn between the behavioral effects of deprivation and the behavioral effects of extinction? Do not, then, deprivation, punishment, and extinction involve the identical process in different forms? The answer to this must be sought in a critical and rigorous examination of the precise stimulus-response contingencies that exist in each case.

A Comparison of Punishment, Deprivation, and Extinction

The withdrawal of a positive reinforcer following a response is punishing; what makes it punishing is the immediate temporal contingency of the two events, namely, of response and withdrawal. In extinction, the withholding of positive reinforcement, or reinforcement in general, is not as sharply discriminable from the standpoint of the organism. Deprivation also involves the withholding of what is presumed to be a positive reinforcement. The effects of these three techniques of withholding differ. In deprivation, the result is activation. Activity of the organism increases after the primary reinforcers have been withdrawn for a period of time. The organism is more likely to behave in ways that produce the positive reinforcer for him. The effect of withholding reinforcement as a punishment is to suppress the response that immediately precedes the punishment. In extinction, withholding results in a fairly gradual decrease in response strength.

Grossly different behavioral effects result from each procedure. Are there, perhaps, more subtle similarities linking them together? Is the extinction procedure in some sense equivalent in its effect to punishment? Both result in decreased frequency of responding. As we pointed out in our discussion of punishment, the effect of punishment is transient. When punishment has been terminated, the organism makes a first tentative response, and finds that it no longer is subjected to suppression. The organism then makes up for lost time. The rate of response increases, and a large amount of the previously

suppressed behavior occurs. Is there evidence that a similar effect is found in extinction? There are reasons to expect that if such an effect were found, it would be small.

However, the logic we have developed would argue that some similarities should exist. In studies concerned with an investigation of properties of nonreinforcement, it has been demonstrated that S^{Δ} (the period of nonreinforcement in discrimination learning) has aversive properties. An organism will terminate a period of S^{Δ} when allowed to do so. In our definition of an aversive stimulus or an aversive state of affairs, it was stated that an aversive state of affairs is one which, other things being equal, the organism will act to terminate. The nonreinforcing condition in discrimination has this aversive character. It has also been shown that employing a so-called "time out" procedure in the training of pigeons will suppress behavior leading to "time out" (Ferster, 1957). "Time out" is a state of affairs where behavior, no matter what it is, goes unreinforced. Finally, in an unpublished experiment (McCrary, 1956), the aversive properties of the extinction procedure were specifically investigated to determine whether there was a compensatory burst of activity following the termination of extinction. As we have just stated, it has been reported (Estes, 1944) that where the organism was punished and punishment was discontinued, the organism compensated with an accelerated response rate. McCrary found that reinforcement of early responses following extinction led to a somewhat higher rate of responding than had been observed in the steady state of reinforced behavior prior to extinction. These data support the prediction that a compensatory spurt should follow extinction, just as it

is known to follow termination of punishment if extinction is similar in its effects to punishment.

As we have seen, aversive and positive stimuli have symmetrically opposite effects with either their presentation or withdrawal. We should note the underlying effect upon the organism resulting from either of the two kinds of motivating conditions. The effects are so immediately similar that they constitute the common sense basis for tension or drive reduction theories in psychology. Extreme deprivation is ultimately damaging to the organism; in fact, it is fatal. At the level of physiology, deprivation represents another aversive technique. Although the two procedures are employed in inverse ways, in reinforcement, their effects are rectified. With positive reinforcement, there is a decrement in the aversive deprivation state of the organism. In negative reinforcement, there is a similar decrement in an aversive state of affairs. Similarities between the two procedures end here because, in positive reinforcement, the decrement in the aversive state of affairs is cumulative. With succeeding positive reinforcers, the decrement successively decreases the state of deprivation to the end that the organism ultimately is satiated. In other words, the reinforcing operation little by little reverses the state of affairs originally established by deprivation. In negative reinforcement, the withdrawal of the aversive state of affairs is a step function. Typically, it is totally reduced and then totally reinstated. Differences in behavior that one observes arising from the two techniques of motivation are easily traced to sharp differences in procedure. In positive reinforcement, the original aversive state of affairs is not fully reinstated following a reinforced response unless the intervals between reinforcement are exceptionally

long. In negative reinforcement, a situation analogous to that which we considered in our treatment of punishment exists. Namely, the experimenter exercises strict control over the contingencies governing withdrawal of the aversive stimulus, but he typically exercises no control over its reintroduction. This cannot help but introduce extraneous controls because the reintroduction of the aversive state of affairs must inevitably punish something. In punishment, where the contingencies of punishment are controlled, there is typically no control over the response or behavioral contingencies existing at the time shock is terminated.

Many of the paradoxical results derived from experiments on negative reinforcement and punishment can be accounted for by this on-off phenomenon. Whether the aversive stimulus is on-off or off-on, the second element in the procedure is usually uncontrolled with respect to the behavior occurring at the time it occurs. Analogous results to those obtained with positive reinforcers might be obtained with negative reinforcers if shock were reduced in intensity by gradual decrements instead of by total termination. At the end of an experimental period, the organism would be returned to its nonmotivated state by having successively reduced by small steps the amount of shock to which it was subjected. This would be analogous to the satiation of an organism that has been surfeited with food pellets. Shock could be built up slowly over time, just as the aversive effects of food deprivation accrue. Similarly, many of the effects of punishment may be more an artifact of the sudden cessation of the aversive state of affairs than of those aversive states of affairs themselves.

In some studies, punishment has been shown to

facilitate learning. If such results were true without complication, they would seriously challenge our entire analysis. Happily, they are explained if we attend to the precise contingencies between behavior and the controlling operations. Learning is produced by the termination of the aversive state of affairs because this, by definition, is negative reinforcement. If a specific action is punished and so suppressed, other behavior takes its place. That other behavior occurs when the aversive state of affairs is terminated. By definition, that other behavior should be strengthened. The fruitful effects of punishment result from termination of punishment, the negative reinforcement of behavior, rather than from the suppression of the punished behavior.

We cannot overemphasize the importance of the delicate temporal contingencies existing between behavior and the environmental operation. It is not sufficient to talk broadly of the effects of punishment; one must examine minutely the precise behavioral and environmental contingencies in a given experiment in order to interpret results. The introduction of any aversive state of affairs serves basically to introduce competing behaviors. We may not notice competing behaviors in connection with motivating operations simply because we do not pay attention to the fact that a sitting or sleeping organism is thereby behaving. When we motivate him, we institute behaviors that compete with resting behavior. When we shock the animal, we initiate respondents and primitive operants that compete with resting action. We wake him up, so to speak.

We have already seen that in some theoretical interpretations of the extinction process, extinction is regarded as the conditioning of competing behaviors. In

extinction, it is difficult to identify what can function as reinforcement. At an operational level, the extinction process is analogous to the original motivating state of affairs in that no systematic consequences follow behavior. Entropy is consequently increased. Previously nonreinforced responses are made more likely to occur. If we look at behavior that previously was reinforced, we see that its probability of occurrence decreases. Probabilities of other behaviors in the situation increase somewhat. Because no particular response is singled out from the rest for reinforcement, this increase is small for a given response class, and therefore may not be immediately evident to casual observation. Certainly other behaviors must take the place of the formerly reinforced response undergoing extinction.

Uses of Aversive Techniques of Control

What can we say, then, concerning the usefulness of aversive techniques of control in practical teaching situations? First of all, we acknowledge that deprivation creates an aversive state of affairs or constitutes an aversive state of affairs for the organism. In this sense all motivation is in some degree aversive. We must recognize that insofar as primary biologically important deficits are concerned, the teacher seldom, if ever, has control over them. A notable exception is governmental use of extreme training measures for purposes of political indoctrination. It is perhaps fortunate that such control is generally out of reach. More common aversive techniques, however, are available to the teacher and these are regularly employed. Threats of failure, threats and

use of penalties involving extra work, threats of and actual use of corporal punishment, etc., are available to the teacher. If one is concerned with only the short-term effectiveness of these techniques in policing the classroom, then punishment is unquestionably effective. No one denies that the use or the threat of punishment will suppress behavior. Presumably, however, no enlightened educator has as his primary objective the policing of his students. This is sometimes a noxious and necessary accompaniment to teaching, but not an educational objective. The teacher who employs punishment can be concerned, moreover, with only the most immediate results of policing because, in the long run, the person so policed will find ways to reduce the aversive state of affairs so that he will come to control the controller.

What if the teacher avoids the stimulus-response contingencies of punishment but employs aversive control in the form of negative reinforcement? The answer is that in order to be terminated, the aversive stimulus must at some point be introduced. It is inevitable with the ordinary aversive control available to the teacher by definition that he punish some behavior or other. This is especially true because once negative reinforcement has been given, the aversive state of affairs must be reinstated before another negative reinforcement is administered. The result is a series of punishments and negative reinforcements. Punishment might possibly not be so marked if the aversive state of affairs could be established gradually, as was stated before, but there is no experimental evidence to support this argument. Also, aversive techniques of control do elicit primitive respondent behavior, which is readily conditioned to any and all stimuli that are consistently paired with it. As stated be-

fore, this results in the punisher taking on conditioned aversive properties that are very slow to extinguish.

Do we then argue for a permissive teaching situation? To this we must reply that the concept of the permissive situation is another of those broad concepts which does not address itself to the particular behavioral environmental contingencies to which it must apply. A truly permissive environment is probably not possible since the physical environment itself sets limits to behavior. The child who falls down inevitably gets hurt. It is furthermore unrealistic from the point of view of the social environment in which the child will live to create such an artificial hothouse environment. An organism trained in such a situation is woefully ill equipped for dealing with the nonpermissive environment of his later years. Realities force us to reject this concept. Limits must be set, and the occasional use of aversive techniques is inevitable. They should, however, only be used when all other techniques have failed.

What, then, is available to the teacher in controlling or motivating students to perform? It has been said that if we eliminate anxiety from the learning situation there will be no motivation. It may very well be, given the particular backgrounds of many children reared in this culture, that there is no other effective motivating device. We may not need to be so pessimistic, however, because there has never been a laboratory animal that was immune to the conditioning of secondary or generalized reinforcers. Ideally, it is the parent's work to establish these as effective controllers of behavior. It is unfortunately true that many parents fail in this responsibility.

The correction of such failures is an arduous and time-consuming task, but it may well be worth the ef-

fort of the teacher, particularly in the lower grades. It is an arduous procedure for two reasons. One reason is that it requires constant attention, at least in the early stages. The teacher must associate himself and his actions with what rudimentary positive reinforcement is available to him. Early failures, of course, begin extinction on its way. There must be objectivity, because the behavior of a child with these deficiencies is unacceptable in the extreme. One must not lose sight of the fact that the teacher's first task is to establish effective generalized reinforcers. Once these are established, he can, from a new base of operations, work to alter the behavior of the child to more acceptable forms. This is not to imply that both processes cannot go on concurrently, but merely to indicate their order of importance. The second reason it is difficult is that obnoxious behavior on the part of a child is aversive to most normal people. The reaction of a teacher, like the reaction of any good laboratory animal, is to act to terminate the aversive state of affairs. Attack, threats, the use of punishment, do have this immediate effect. Punishment is more immediately reinforcing to the punisher than almost any other technique, because it so quickly and effectively terminates what is aversive to the punisher. Once the teacher recognizes this tendency, he can compensate for it. Few people are born with the patience and objectivity required for this task, but such virtues are laboratory skills that can be acquired. We leave the teacher, then, with the injunction to work insofar as possible with generalized reinforcers that have been established by the general social environment in which the child exists. Where possible, the teacher should use those particular generalized

reinforcers that are peculiar to the child's cultural, ethnic or familial background when the teacher knows that background and has access to appropriate reinforcers.

Traditional Concepts of Motivation

What then of motivation? It is clear from the foregoing discussion that we have said very little about the classical area of motivation. In fact, the original plan for the present analysis did not include a treatment of motivation per se because topics included in the classical discussion of motivation cut across many different areas of analysis that are more effectively dealt with in other terms. A typical introductory textbook in general psychology includes under the heading of motivation discussions of the homeostatic mechanisms of the body, the endocrine system, stomach contractions, temperature regulation, the female sex cycle, the maternal behavior of rats, sleep, the effects of sensory deprivation, secondary and generalized reinforcement. It is easy to see why the educator finds little in such writings to help him stimulate his students to learn. With the exception of the discussion of secondary and generalized reinforcers, there is little that tells the teacher what he can do to manipulate the variables involved in "motivation." Certainly the most he can do with regard to the female sex cycle in his students is to ignore it and to hope for the best.

What, one might ask, is the place of the traditional concepts of motivation such as incentives and goals? What about acquired drives and motives? Are instincts not relevant to an analysis of motivation? In our present

treatment, incentives and goals are treated, as is the concept of reward, under the category of reinforcement. Incentives are reinforcers. Goals are the reinforcers or the reinforcing state of affairs that reasonably may be expected to follow accomplishment. Unfortunately, the concept of *goal* carries with it philosophical connotations of purposiveness that for a scientific analysis are objectionable. It is a way, however, of talking about events that do control behavior through reinforcement. We shall not embroil ourselves in a detailed discussion of teleology; let it suffice to say that an account of a present action in terms of a future event is not permissible. It is for this reason that no serious place is given this concept in a scientific analysis of behavior. It is another of the many prescientific ways of explaining human action which are in the process of being discarded. No one wishes to deny that there is direction in human action. We do not deny that behavior has characteristics that may be named *purposive*. We do reject the naming, since we must reject the metaphysical interpretation of the causes of behavior such naming represents.

We have already seen that explanations of behavior and the learning process in terms of drives and drive reduction are inadequate. The drives with which we concerned ourselves in Chapter Three were basic, primary, biological drives. Many contemporary theoretical behavior systems are built upon concepts of acquired drives and motives. Do the same objections apply to the use of these concepts as apply to primary drives in explaining behavior and learning? The answer must be yes. Moreover, objections to the hypothetical nature of basic drives are even more pointed when raised against acquired drives, for acquired drives are even farther removed from

empirical manipulation and observation than are primary drives. By what deprivation operation does one increase the "drive to excel"? By what procedure is this drive satiated? It is sadly true that recourse is often taken to such concepts to provide fictional replies to questions about the causes of behavior—replies that stop questioning, but which do not *answer* questions.

We should be less than honest, however, if we did not point out that the present analysis deals with the same phenomena as those dealt with in terms of acquired drives and motives. The emphasis, however, is shifted. The shift of emphasis is not trivial; the key to the power of this analysis lies exactly in this shift. We have concerned ourselves with primary reinforcers as controllers of behavior. We have not become involved in theoretical issues regarding inferred states of the organism relative to such reinforcers and their effects. An explanation of reinforcement in terms of changes in an inferred drive state does not increase our ability to use reinforcers to control behavior. Similarly, we have concerned ourselves with secondary and generalized reinforcers as controllers of behavior. We know how to establish such secondary or conditioned controllers. A hypothetical explanation in terms of learned drives does nothing either to add to our real knowledge of what is going on or to increase our ability to use these reinforcers to control behavior. It is more likely that explanations in terms of such constructs will embroil us in fruitless debates of theoretical dogma.

Finally, fictional explanations of behavior in terms of learned drives must be rejected on the same grounds that explanations of behavior in terms of instincts have been rejected. As a hypothetical construct or intervening variable, the concept of acquired drive has a legiti-

mate place in scientific analysis. As an *ad hoc* explanatory device, it has no place. To say simply that a pupil performs at a high level because he has a drive to succeed or an instinct to excel is to say no more than that a chicken crosses the road to get to the other side. Both explanations deserve exactly the same respect as serious accounts of behavior. Instincts and learned drives provide easy verbalisms for explaining away behavior; they both are vulnerable to multiplication as there are no rules for their use under such circumstances. Without careful control, one may discover that there are drives for just about everything. Eventually, it is evident that naming a new drive tells us nothing about what we can do with it. We shall dismiss *instincts* with a definition. Psychologists are agreed that for a phenomenon to be classed as instinctive it must meet three criteria. It must be a complex pattern of behavior; that is, a simple knee jerk reflex does not qualify. It must be unlearned. It must be universally present in the species. If these three criteria are met, it is clear that very little in human conduct qualifies as instinctive. One exception invalidates inclusion. It is just possible that learning itself might qualify. Complex modes of human action, however, do not.

One point in the foregoing discussion may concern the careful reader. We saw in the chapter on reinforcement that secondary or generalized reinforcers are established through their association with primary reinforcers. One may ask whether it is not necessary then that the teacher have control, at least for a limited time, over the primary reinforcers of the individual's behavior in order that he may establish generalized reinforcement. If this is true, then the prospects for establishing effective generalized reinforcers are bleak indeed.

There is, however, a possible alternative explanation of the establishment of secondary reinforcers which has not, to this writer's knowledge, been investigated experimentally. This explanation derives from the Guthrian contiguity theory of learning. Guthrie says that a significant change in the environment that reliably follows a member of the desired response class produces those changes which we say represent reinforcement. How does this relate to what we know of primary and secondary reinforcers? Primary reinforcers by their physical nature have the required immediate and reliable effect. Social reinforcers are not as reliable. They generally are more reliable in the small child's immediate family environment, and this is possibly the reason that secondary and generalized reinforcers are usually established in the family setting. It is possible that the establishment of secondary reinforcers through association with primary reinforcers is an artifact of the more necessary condition of the reliability of the behavioral consequence. In experimentation with infrahuman animals, things that are associated with primary reinforcers such as the sound of a food magazine or the flashing of a light, also occur with high regularity. It would seem a worth-while experiment to determine whether it is the pairing with the primary reinforcer that is necessary for the establishment of secondary reinforcer or whether it is simply the consistent reliable consequence of action that establishes a neutral event as a secondary reinforcer. If the latter were true, it would have significant implications for the theory of learning in general. In practical fact, this is what is available to the teacher in most instances as reinforcement.

No discussion of motivation is complete without a

consideration of the contributions of Sigmund Freud. The Freudians concern themselves almost exclusively with sex as the primary motivator of human behavior. Sex is, of course, broadly defined; but even so, it remains sex. Freud's concern with motivation was in its unconscious functioning. The analysis of behavioral difficulty can, according to Freud, best be undertaken in terms of those variables of whose existence the person is usually unaware. This systematization may be useful to the clinician, but it is not particularly helpful to the teacher. Concerning oneself with the unconscious motivations of the students may do more damage to the student-teacher relationship and create more havoc with the learning process than it does good. From our earlier discussion of the effects of aversive control, we know that aversive control leads to the suppression of behavior. Where those effects lead to suppression that is so complete that the individual is "unaware," where motivation, in the Freudian sense, is unconscious, there are excellent reasons from the point of view of the organism's economy why the organism should be unaware. Tampering with such variables and such behaviors is not the proper province of anyone not specifically trained and professionally responsible to do so. Not only does concern with the unconscious motivations of the student not provide a particularly useful motivational handle for the teacher, but, in effect, it directs the teacher to a kind of panic button that he should take special care to leave alone. This is not to say that the teacher has no legitimate concern with the emotional problems of his students as they relate to classroom performance, or that he should not bring significant emotional problems as he encounters them to the attention of qualified authorities. Nor does

this mean that the teacher has no legitimate business in learning enough about his student to determine factors that are effective, or might become effective, as reinforcers. It is simply a caution against the perhaps well-intentioned but extremely dangerous practice of bringing amateur psychoanalysis into the classroom. It is also too often true that pseudoprofessional rationalizations of a student's failure to learn in terms of the "psychic" life of the student are a smoke screen for professional incompetence. Skinner (1954) has summarized what the teacher has available for the control of behavior.

> In the first place, what reinforcements are available? What does the school have in its possession that will reinforce a child? We may look first to the material to be learned, for it is possible that this will provide considerable automatic reinforcement. Children play for hours with mechanical toys, paint, scissors, paper, noise makers, puzzles. In short, with almost anything which feeds back significant changes in the environment and is reasonably free of aversive properties. The sheer control of nature is itself reinforcing. This effect is not evident in the modern school, because it is masked by the emotional responses that are generated by aversive control. It is true that automatic reinforcement from manipulation of the environment is probably only a mild reinforcer and may need to be carefully husbanded, but one of the most striking principles to emerge from the recent research is that the net amount of reinforcement is of little significance. A very slight reinforcement may be tremendously effective in controlling behavior, if it is wisely used.

CHAPTER FIVE

Complex Processes

To speak of complex processes is not to imply a discontinuity between the basic processes which we have already considered and the more elaborated forms of behavior. By complex processes, we refer simply to the processes relating those behaviors to a more complex set of variables. Historically, writings on complex processes have dealt with the so-called "cognitive" aspects of behavior. Implicit in these treatments has been the assumption that the organism shifts gears, so to speak, and that the essential nature of these so-called higher activities is different from the simpler behaviors of the organism. Our position necessarily is that no such discontinuity exists. It is incumbent upon us to adduce evidence in support of this claim, for, in effect, we claim that the behavioral processes involved in the conditioning of a bar press in the rat are essentially the same as the processes underlying speech and thought in the human. A series of researches are relevant to this position.

The first is that of Greenspoon (Greenspoon, 1955). He demonstrated that one can condition the emission of verbal operants with mild generalized reinforcers. Sub-

jects of this experiment were asked to say words, singly, without repeating, without making statements or phrases, and without counting. He categorized the verbal responses of his subjects into those which were plural nouns and those which were not. Reinforcement consisted of the experimenter's saying "Mm hmm" following the emission of a plural noun. It was demonstrated that with this form of reinforcement the frequency of plural nouns increased during conditioning. Verplanck reports a series of experiments wherein the experimenters established control over more complex forms of verbal behavior. In one study he demonstrated that simple operants, such as raising the hand, can be conditioned in what amounts to a game reminiscent of "Twenty Questions" (Verplanck, 1956). Later, he and student experimenters succeeded in conditioning the frequency of occurrence of statements of opinion (Verplanck, 1955). In these studies, the experimenters used agreement with the opinion statements that the subjects made as reinforcers in some groups; in other groups, a paraphrasing of the opinion statements were used as reinforcers. The effect in both cases was an increase in frequency of opinion statements.

Indirect evidence that something akin to the processes we have considered is operative in the development of language is provided by phonetic analysis of language in young children (Irwin, 1947). In a child 10 months of age, the consonant phoneme with the highest frequency of occurrence—41.29 percent—is that represented by the letter *H*. In adult speech, this consonant phoneme accounts for only 2.66 percent of the verbal production of an English-speaking adult. The high percentage drops gradually in the child, until at the age of 30 months, this particular phoneme accounts for only

7.65 percent of his total output. In adult speech, the consonant with the highest frequency of occurrence, 11.85 percent, is represented by the letter *N*. This particular consonant phoneme accounts for only 1.03 percent of the verbal production of a 10-month-old child. Comparing the verbal output of the 10-month-old with that of the 30-month-old child, we see that by the time the child has reached nearly 3 years of age, this particular consonant accounts for 9.49 percent, which is hardly significantly different from the adult frequency.

These changes in the distribution of verbal production are not simply the result of maturation; they are produced by the mediation of the appropriate verbal community. The verbal community reinforces that verbal behavior most nearly approximating its own. These percentages would, of course, differ in a non-English-speaking sample, but the trend of approximation of infantile speech to adult speech would naturally occur. A large and valid body of experimental evidence exists to support the contention that not only are the processes of conditioning of simple actions continuous with the processes of conditioning of more complex actions, but also that the processes of conditioning do not differ in their essential nature from species to species. We need not make apology in defense of commonly accepted practice in biological science. It is legitimate to extrapolate knowledge gained from studying the processes of one species to the study of processes in another.

Let us now turn our attention to the complications of the learning process and their effects, which give rise to the varied and intricate behaviors observed in the human organism. It is our position that discrimination learning is the basis upon which all of the so-called com-

plex behaviors are constructed. Special motor skills also require response differentiation, but the establishment of the subtle environmental control over behavior that defines the meaning of complex action rests upon discrimination training.

Stimulus Control

The concept of learning as it applies to education may occasionally refer to simple changes in response probability or to the acquisition of motor skills. But generally, learning refers to stimulus control through discrimination training. When we say that a person has learned something we do not usually mean that he is seen to exhibit a certain class of behavior with greater frequency, but rather that his behavior has been brought under sharp environmental control. This is even true when we consider the acquisition of skills such as playing the piano. A person who plays the piano well does not simply play more notes more frequently. The most important change in his behavior is the fact that it is now under the control of printed music. The behavior of the accomplished pianist must be brought under the control of such external stimuli at some point. The behavior of the concert pianist moves beyond this stage to a level where his playing is controlled by sequences of his own behavior and by the auditory stimulus provided by his playing and the music of the accompanying orchestra. In all cases, control of the response by independent external stimuli is essential.

That some type of stimulus control is necessary in most learning is evident when we consider performances

that require pacing in time. Here the application of a rate or frequency measure as our criterion of learning might be inappropriate. Some kind of index might be used to indicate the degree to which discrimination learning has progressed—an index combining the behaviors observed and their appropriate and inappropriate circumstances. Such an index has been reported (Green, Sanders, Squier, 1959).

By the time the child enters school, he has been subjected to the major part of the differentiation training demanded by his culture. Most of the motor skills that are necessary for his effective interaction with his environment have been established. The way in which the child's verbal performance comes to approximate that of the adult exemplifies the process of response differentiation. Insofar as experimental procedure is concerned, response differentiation involves a different procedure from stimulus discrimination. The experimenter uses the so-called method of approximation in shaping the behavior of the organism. He observes on-going behavior and selects some portion of that behavior that approximates to some degree the end product he seeks to establish. The experimenter then reinforces that approximation. He reinforces it on its next occurrence. Over succeeding instances, he gradually demands that the behavior of the organism more and more closely correspond to the finished product, until finally it does correspond exactly. The experimenter makes reinforcement contingent upon the organism's having emitted some form of response that approximates in more and more precise degree the final repertory that is to be established. From the subject organism's point of view, differentiation *is* discrimination. For the subject, the

stimuli that are made significant for him by this procedure, the stimuli that are ultimately placed in control of behavior, are those stimuli produced by his own behavior and its effect upon the environment. The subject is required to discriminate stimulus feedback from its own actions. This feedback may be of two types: the immediate kinesthetic feedback of muscular action, and the consequences of locomotion within the environment. Conditioning the organism to approach more closely a particular section of the environment makes the stimuli associated with that portion of the environment more salient to the organism, and increases the ease with which that portion of the environment can control behavior. This particular argument that response differentiation is a special class of stimulus discrimination will be important in our later consideration of the way in which programmed instruction operates.

Most of the education process involves the establishment of more and more highly integrated forms of stimulus control. It is appropriate that we examine in some detail the variables of which such discrimination is a function. Let us consider first the paradigm of discrimination learning. In discrimination training, the experimenter makes reinforcement of a response contingent upon the presence or absence of particular aspects of the environment. Certain stimuli when present are considered to be appropriate occasions for the behavior, and behavior is reinforced when it occurs in the presence of those stimuli. Should behavior occur in the absence of such stimuli it goes unreinforced. Technically speaking, we refer to the occasions for reinforcement as S^D. We refer to those occasions in which behavior goes unreinforced as S^Δ. As a consequence of this differential rein-

forcement, behavior comes to be emitted in the presence of S^D and extinguished in the presence of S^Δ. A sharp difference in responding to the two stimulus contingencies is established. S^D comes to evoke the operant response, but not with the high degree of reliability with which respondent behavior can be called forth. Respondent behavior is elicited by its unconditioned stimulus, but the control of an operant by an S^D is not nearly so precise, even though it may superficially appear to be similar. The properties of the environment to which discrimination is relevant may be so subtle as to make concrete reference to them as stimuli often unrealistic. For that reason, it is suggested that S^D and S^Δ should be conceived of more as *occasioners* for behavior than as stimuli. The relevant properties of the environment upon which discrimination is based may be quite obscure. Discrimination may be of such things as color of objects, shape of objects, number of objects, or combinations of these properties.

At one point Skinner and Ferster considered dedicating their book *Schedules of Reinforcement* "To the mathematicians, statisticians, and scientific methodologists with whose help this book would never have been written" (Skinner, 1958). Insofar as the present mathematical developments in theory are inappropriate to or inadequate in accounting for commonly known laboratory phenomena, there can be no quarrel with this rather flippant observation. However, we shall not hesitate to make use of certain analyses that have been developed in the context of mathematical learning theory where these may be appropriate and do no violence to the facts. We shall employ these concepts as helpful analogies, as a kind of shorthand in describing behav-

ioral events to which they may apply. Particularly helpful is the analysis of the stimulus situation or context within which behavior occurs in terms of statistical learning theory (Estes, 1950). Equally helpful has been the incisive and rigorous definition of relevant parameters within the learning situation, such as the initial state of behavior of an organism. The initial probability of a response, $\overline{P}_0$, is a shorter symbol for the operant level of a response. The adoption of a probabilistic shorthand need not imply the adoption of the theoretical system from which it is drawn. The sole justification for such an adoption is to make communication more convenient. This is our intention.

As was pointed out in the discussion of the process of extinction, the environment is in a state of continuous flux. In the earlier discussion we spoke of stimulus elements. Operationally, stimulus elements are defined in terms of certain properties of the learning curve. They have no independent existence beyond this. To test a theory, we may equate certain aspects of the physical environment with these hypothetical objects and perform experiments with reference to them. Experiment then shows or does not show correspondence between actual behavioral processes and the hypothetical processes of theory. The experimenter can employ conditioning techniques at his disposal to make various aspects of the stimulus field effective as stimulus elements. Arguments in the experimental literature with respect to the question of whether a stimulus element is a concrete property of a stimulus or whether the relevant aspect of a stimulus is a pattern of such components must now be viewed with the knowledge that what is relevant to the organism is that which is dictated by the controlling

variables in a given setting. In one setting the component may be relevant as a stimulus element. In another setting, patternings of elements may comprise components of a larger order of magnitude. It is not necessary to choose between elementaristic or pattern conceptions of the physical environment; both have relevance to behavior. Their special relevance can be understood only in terms of the specific context in which they are studied. It is to be hoped that future literature on the learning process in an educational setting will not be composed of recapitulations of old experiments dressed in different methodological robes. Much of what has gone before is instructive as a lesson in what should be avoided in scientific inquiry.

Stimulus Elements

Bearing in mind, then, that the particular properties of the stimulus environment that are to form the basis of discrimination are selected and determined by the experimenter, we may refer to these properties or combinations of these properties as stimulus elements. We do not intend to imply anything concerning the theoretical nature of conditioning or discrimination by the use of the term, but rather employ it for convenience. The stimulus elements in any discrimination problem belong to any one of three subsets. The first is composed of those elements which are present exclusively as occasions for reinforcement. The second of those which are never present when reinforcement is given. The third of those which are sometimes present when reinforcement is given and sometimes present when the occasion for nonrein-

forcement obtains. It can be argued that even simple conditioning or extinction inevitably creates a discrimination experiment.

As discrimination is classically studied, a single response class is investigated as a function of one of two stimulus sets presented either simultaneously or in temporal sequence. We do not need to confine ourselves, however, to a single response class; nor need the stimulus field be categorized into but two such sets. The general case involves a mapping of the response space as it relates to *N* overlapping stimulus subsets. Although it is possible to conceive theoretically of such complexly interrelated analysis, time would prohibit experimental investigations of such order of complexity. Moreover, it remains true that regardless of the number of response classes which the organism is engaged in emitting, they may always logically be subdivided into two classes, classes of Response A and classes of Response $\bar{A}$. For these reasons we shall confine our attention to the simplified situation.

To return to our analysis of the stimulus field in discrimination, we see that those elements, those properties of the environment composing S^D, are of two types. The first are those elements upon which discrimination is based. These are the defining properties of the concept, so to speak; they are the elements that are reliably present and that actually constitute the occasions for reinforcement of the appropriate behavior. The second category of elements, which we shall call the intercept elements, are those which are present both in S^D and S^Δ. Symmetrically, S^Δ is composed of two types of elements: those which are reliably associated with conditions of nonreinforcement, and the intercept element which S^Δ

shares with S^D. The intercept elements have long been recognized in classical studies of learning as extraneous variables, uncontrolled stimulation, and more recently, in the context of information theory, as "noise." To sharpen discrimination, it is desirable to reduce the contribution of the intercept elements relative to the nonintercept elements.

Variables in Discrimination Learning

Since our position is that the complex learning with which programmed instruction is concerned is a form of discrimination learning, let us elucidate those procedures which contribute to the efficiency of learning. What are those variables which most rapidly and effectively establish a behavioral repertory and develop appropriate stimulus control? A continuing series of researches exploring the parameters of an operant discrimination of visual material in human subjects, provides relevant information (Green, 1955, 1956, 1957, 1958). The relative contribution of intercept elements has been effectively shown to govern not only the rate at which discrimination learning proceeds but also terminal or asymptotic performance. Some evidence exists to indicate that with the proper weighting of elements within the three stimulus subsets, subjects may more rapidly learn a discrimination to a certain level of excellence; but the level of excellence they attain is lower than is attained with other combinations of the weighting of elements. These latter combinations of elements at the same time give rise to more gradual conditioning.

The concept of the *difficulty* of a discrimination task becomes ambiguous in the light of such results.

A second variable important in the efficiency of discrimination learning is the ability level of the students. Theoretically, all students could master all the discriminations presented to them if the slower students were given unlimited time to learn. Practically, of course, the slow or retarded student does not have unlimited time in which to learn. At some point terminating the educational process, he is found to be short in his accomplishment as contrasted to the brighter student. Theoretically it can be shown that even if bright and dull students were allowed to reach the same criterion of excellence eventually, important differences would exist between their behaviors. While the mean performance which they would attain might be equal, the variability of their performances would differ. The behavior of the slower student should be more variable than that of the faster student. Such indeed is the case, and not only in a simple visual discrimination (Green, 1961a); it has also recently been found to be true in the more complex behavior of students responding to programmed instruction.

Another important variable in the efficient conditioning of discrimination is the constancy of the stimuli to which the individual must respond. If elements are constantly moving into and out of the available stimulus matrix, discrimination learning takes longer than if the elements composing the discrimination remain stable. If discrimination is based upon a small number of elements, then it is most efficient to present these elements with minimal variation. As we have seen in considering extinction, there may be advantages in increasing the variability of stimulus elements. Such advantages lie in

increasing the generality of discrimination and in increasing the resistance of the discrimination to extinction. The gain in generality and resistance to extinction might outweigh the disadvantage in the slower rate of training. Increasing the variability of S^D increases the relative size of S^D, since the total stimulus set consists of all those elements which have been employed over a period of time.

Germane to our discussion of the functions of stimulus elements in discrimination training is a controversy that has developed among the researchers studying discrimination from the standpoint of mathematical learning theory. This is the question whether the subject forms his discrimination on the basis of discrete elements or components of the stimulus situation, or whether he forms his discrimination on the basis of patterns of such components. Our resolution of this controversy is pragmatic. The experimental subject forms his discrimination on whatever basis the experimenter contrives that he should solve it. If the experimenter contrives a situation where the most effective solution of the discrimination problem is in terms of components, then the subject will learn to discriminate components. If the most effective solution is in terms of patterns of elements, then the subject will base his discrimination upon patterns of elements. It is possible that there is a priority of solutions, given the possibility that the subject can solve the problem on either basis. There is some evidence to indicate that the subject initially begins to discriminate on the basis of elements and then switches at some point in the process to patterns. This is still an experimental question and one suspects, given the flexibility of experimental subjects and the ingenuity of experimenters, that any

such priority can be reversed by appropriate manipulations.

Another relevant aspect of the training situation is the interval between training sessions. In effect, this boils down to the effects of differences in the relative size or duration of S^D as against S^Δ. Since discrimination involves not only the conditioning of the response to the appropriate stimuli but also the extinction of that response to another set of stimuli, one might argue that the optimal discrimination training situation should present S^D and S^Δ in equal proportions. If all of the behavior of a student is reinforced in the presence of one aspect of his environment, he may be thoroughly conditioned to respond in the presence of that portion of his environment, but he is also likely to generalize to other parts of his environment. He has not been made aware that there are important parts of the stimulus field to which such behavior is inappropriate. Also, if the conditioning history of the student is composed almost exclusively of extinction in the presence of S^Δ occasioners, the response will never come to exist in any considerable strength.

In Figure 6 it is seen that where the number of S^D occasioners equals the number of S^Δ occasioners ($\pi = 0.50$), the most sharply accelerated rate of responding is produced in S^D, and the highest over-all terminal rate of responding is obtained as long as the experiment lasted. The second most rapid acceleration in S^D was produced where one-fourth of the discriminative stimuli in conditioning were S^D ($\pi = 0.25$); the lowest was produced where three-quarters of the stimuli in the conditioning series were S^D ($\pi = 0.75$). Although the two latter curves are the reverse of what might be reasonably expected, it is nonetheless true that both conditions where there was

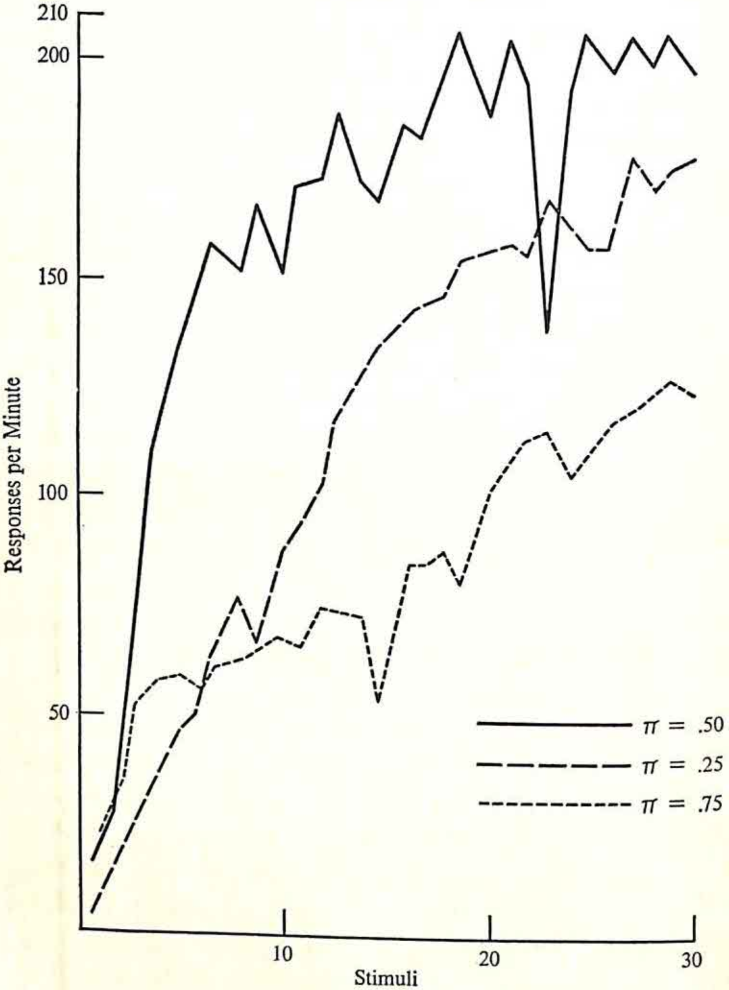

FIGURE 6. S^D curves in discrimination with differing proportions (π) of S^D occasioners in stimulus series.

an imbalance between the proportion of S^D and S^Δ in the stimulus series produced inferior discriminations in the context of this experiment, thus supporting our hypothesis. No detectable effect upon performance in S^Δ is seen to result from any of these three conditions.

In the experimental study of discrimination learning, there are technical problems that are of no practical concern to the educator. We shall discuss them, however, because they may become relevant as the techniques of programming become more highly developed. One must not use simple alternation of S^D and S^Δ, because one property of the environment upon which the organism is capable of basing discrimination is temporal sequence itself. For this reason, the sequence is determined in a random manner. Another concern is that the intervals of S^D and S^Δ not be too short. This is particularly important where reinforcement is based on a schedule (Green, Sanders, Squier, 1959).

Schedules of Reinforcement

Perhaps the most significant set of variables controlling the speed with which discrimination is established and controlling the viability of discrimination has to do with schedules of reinforcement. Conditioning a rat to press a lever is usually done with continuous reinforcement (CRF). Every response is reinforced. This produces the most rapid conditioning, but, once reinforcement is withheld, CRF also results in the most rapid extinction. The contingencies between response and reinforcement can be manipulated in two basic ways, giving rise to four general types of schedules. Each schedule produces char-

acteristic forms of responding. At least three of these schedules are immediately relevant to education, and all of them may become significant in the developing technology of programmed instruction. Reinforcement may be given either on the basis of the amount of behavior emitted or upon the passage of time. Moreover, reinforcement may be programmed in a regular or an irregular manner.

In ratio reinforcement, reinforcement is contingent upon the emission of a certain number of responses. Where reinforcement is made contingent upon the repeated emission of a fixed number of responses, the technique is called fixed ratio reinforcement. Every *N*th response is reinforced. The ratio refers to the number of responses per reinforcement. The irregular ratio schedule is called variable ratio reinforcement; here every *N*th response *on the average* is reinforced. Where reinforcement is primarily based upon the lapse of time, we speak of interval schedules. In interval schedules of reinforcement, the first response following an elapsed period of time is reinforced. Reinforcement is not totally independent of behavior even here, for if the reinforcement should be delivered independent of the particular response, an effect which has been called experimental "superstition" is produced (Skinner, 1948a).

When reinforcement is given following the first response after a constant interval, we speak of fixed interval reinforcement, as we refer to fixed ratio reinforcement where the ratio is not allowed to vary from reinforcement to reinforcement. In fixed interval reinforcement, as in fixed ratio, the intervals are regular. The first response after every 30 seconds or every 5 minutes, for example, is reinforced. Timing of the first and subse-

quent intervals starts from the beginning of the experiment and is independent of whether a response is made within that particular interval, at the end of that interval, or at the end of a succeeding interval. In the conventional interval schedule, once the specified interval has elapsed the organism is eligible for reinforcement, no matter how long he may delay. Variable interval schedules are constructed similarly to variable ratio schedules. The first response following an interval of an *average* duration is reinforced.

That interval schedules have effects other than those of ratio schedules is well known, but the implication of this in the analysis of basic conditioning processes is often overlooked. An illustration is found in a demonstration directed at the response unit hypothesis mentioned in Chapter Three. This demonstration was performed by students in an introductory laboratory course at Dartmouth College. Reinforcing contingencies were set so that one pigeon was reinforced, say, for every 50 responses on a fixed ratio schedule of reinforcement; another bird emitted 50 responses on a fixed interval of reinforcement, and was also reinforced on the 50th response. According to the response unit hypothesis, the effect upon response strength should be the same for the two animals. The same amount of work was required under both conditions. However, resistance to extinction under the fixed ratio condition was much greater than was resistance to extinction under the fixed interval condition—even though the units of work required per reinforcement were equal. Over-all rates of responding during the conditioning were also not equivalent. Fixed ratio reinforcement produced much higher rates than fixed interval. This demonstration again points up the impor-

tance of momentary response-reinforcement relationships as they shape the over-all pattern of behavior.

To return to the specific interval schedules, the schedule with intervals of constant length is called fixed interval reinforcement; the schedule in which intervals vary around some mean value is the variable interval schedule of reinforcement. Studies of schedules of reinforcement as applied to discrimination learning have shown that the sharpest discriminations are generally established with ratio schedules. The highest rates of responding in S^D are produced by ratio schedules, but discriminations are established more rapidly as the values of the schedules employed more closely approximate CRF. In one-hour experimental sessions, the precision of discriminations established under interval reinforcement is markedly inferior to the precision of discrimination established under ratio schedules. The most effective discriminations are obtained with variable ratio of reinforcement.

The effects of schedules of reinforcement is itself a large area of study. It is inappropriate to attempt a comprehensive review of the relation of behavior to these variables here. The reader is referred to Ferster and Skinner (Ferster and Skinner, 1957) and to other primary researches in the field. It is not out of place, however, to call attention to certain behaviors of students that are traceable to the peculiar operations of schedules of reinforcement. One example is that of the student whose behavior is controlled by an interval schedule. He studies little, if at all, immediately following completion of an assignment or examination. He begins to increase the frequency or intensity of his studying only as the time

for the next assignment or examination approaches. Observation of his behavior reveals that he increases his studying frequency to a maximum immediately preceding the due time of the assignment. Reinforcement here is, as usual, the negative reinforcement derived from successful completion of the assignment. Successful completion may have positive reinforcing properties, but the salient feature is the schedules reduction in the aversive threat of failure. That this is the primary variable is evidenced by the fact that the temporal scalloping characteristic of fixed interval reinforcement develops. If the positive reinforcers implicit in completion of the task were of primary importance, the scalloping would not appear, because the positive reinforcement is always potentially available.

Fixed ratio reinforcement produces high rates of responding. It is exemplified in the control of human behavior through piecework schedules. It is known to generate high rates of responding that can be detrimental to the organism. For this reason, legislation has been passed outlawing this practice in industry. Fixed ratio reinforcement produces a higher resistance to extinction than does CRF. The performance of the organism in extinction following fixed ratio reinforcement is characterized by high rates of response broken intermittently by cessation of activity. As extinction progresses, the breaks become more and more frequent and of longer duration in contrast to the rather smooth, negatively accelerated curve of extinction following CRF and some interval schedules. The effects of this schedule are manifest in the behavior of the student who simply cannot write another word on his term paper, although he can write a ten-page letter

to his girl friend. When he does succeed in writing his assignment, he is likely to do so in one frantic burst of activity.

Most social reinforcers are programmed on variable ratio reinforcement, and all gambling devices employ the contingencies of this schedule of reinforcement in their payoff systems. It has been said that nature reinforces the efforts of scientists on a schedule of this type. The effect of this schedule is to produce, for comparable values, even higher rates of activity than are found with fixed ratio reinforcement. Moreover, the resistance to extinction of behavior conditioned on this schedule is higher. The third schedule is fixed interval. We have noted that temporal discriminations are formed in this schedule, producing the so-called scallop effect. This results from the fact that the organism comes to discriminate the fact that once his behavior has been reinforced, it is unlikely that this behavior will be reinforced again very soon. It is not possible for him to get reinforcement successively within very short intervals. The result is that immediately following reinforcement, the organism ceases to respond and only gradually does his rate of response accelerate toward the end of the interval.

In all schedules of reinforcement, resistance to extinction is higher than in CRF. This is true of fixed interval schedules also. The scalloping effect carries over from conditioning under fixed interval reinforcement to the extinction period, so that the extinction curve in FI is characterized by the temporal patterning of responses characteristic of conditioning. In variable interval reinforcement, the rate of behavior has a low rate of responding, as is the case in fixed interval reinforcement. Unlike FI, however, this schedule provides a very smooth, steady

state of behavior that is admirably suited as an experimental base line against which the effects of other parameters, such as drugs, may be assessed.

A major problem in classroom teaching is attention. Particularly in the lower grades, children's span of attention is short. We regard attention as the control exercised over behavior by a particular stimulus. The classical view of attention span is that it is a characteristic of the child that is relatively immutable, and presumably grows or increases as a function of age, following some sort of maturational process. Our view is that the attention span is subject to manipulation by the same types of environmental variables as other behaviors. The problem of attention emerges in different contexts. Holland (Holland, 1958) studied vigilance in connection with radar monitoring and showed that vigilance or attention span could be altered as a function of various schedules of reinforcement. The problem of attention in small children was directly investigated in a study designed by R. T. Green and subsequently reported as a part of a larger study (Bullock and Maline, 1958). In the latter study, a technique was developed that could be used to study attention span as a function of other variables. Attention span, like other concepts encountered in educational theory such as reading readiness, is an instance of the rich possibilities open to experimental research in education. Many such concepts constitute problem areas that hopefully might yield to experimental analysis when laboratory techniques are brought to bear upon them. With respect to reading readiness, Moore has shown that children can successfully be taught to read at the age of 3 years using automated techniques of training (Moore and Anderson, 1959).

Chaining

Finally, we should consider a procedure that in many treatments of the learning process is included as a separate basic process. This is the procedure of *chaining*. By chaining the organism is taught to produce, by means of his own action, discriminative stimuli which, in turn, evoke other behaviors on his part that eventually lead to reinforcement. Chains of behavior can be quite elaborate and composed of many steps. It is under the general category of chaining that we include studies of serial learning. The traditional memory drum experiment employing nonsense syllables falls under this heading. We do not regard this kind of verbal learning as being apart from the conditioning of verbal operants. It has been suggested that learning in such a setting, in some sense, fits a different paradigm from that of simple operant conditioning. Zeaman has argued that the write-in type of teaching machine is modeled after the memory drum (Zeaman, 1959).

We shall consider in detail the properties of various types of machines in a later chapter. For the moment, let us note that the write-in machine involves the visual presentation of material in a sequence. The student makes a response to a frame or item of material either by completing a statement or by answering a question. He makes his response by writing his answer on a tape. In an early model of the write-in machine, those items which had been missed by the student were presented to him again. In this regard, the write-in machine in some respect resembles a memory drum. Let us note, however, the precise task required of an experimental subject fac-

ing a memory drum. The drum presents a fixed number of items in an unchanging serial order. The sequence does not vary from trial to trial. In the customary method, the anticipation method, when the starting symbol or any of the succeeding items in the series is presented, the subject is to anticipate what the next item will be in the series. When the subject can anticipate correctly the entire sequence of items for a predetermined number of trials, he has achieved whatever criterion of learning has been set. This procedure differs in several significant methodological ways from procedures used with the write-in machine. The type of fixed stimulus control, the significant effects of sequence, etc., are not important features of the write-in machine. We do not argue that the two devices require different processes of conditioning for the establishment of the repertories demanded by each device. Zeaman argues that the main difference between the memory drum serial learning procedure and the write-in device is in the self-pacing feature of the latter. We argue that the memory drum and the procedures employed in the conventional study of serial learning represent special limited cases of conventional operant conditioning. They represent cases where extreme temporal and environmental limits are imposed upon the classes of behavior to be conditioned. Moreover, the severe restrictions so imposed produce in that behavior rather unusual artifacts—artifacts to which much of the serial learning literature is devoted. Much effort has been expended in these studies to develop control techniques to overcome these artifacts.

This writer is well aware of the importance of the phenomena of serial learning as they have determined the nature of psychological learning theory for the past

few decades. One wonders, however, whether more fruitful theorizing might have been accomplished had not so much emphasis been placed upon and so much material been drawn from a rather peculiar training technique. Zeaman suggests that there is much in the older literature on serial learning which has direct applicability to programmed instruction. I would raise two objections to this. Even if a great deal of relevant information does exist there, the problems of translation of technique and concepts into practical use are formidable. I also would argue that the technique of itself might apply to certain special forms of programmed instruction but that it in no sense represents the general case.

CHAPTER SIX

The Concept of Programmed Instruction

We take the view that the basic paradigm of programmed instruction is that of the interaction of two persons. The basic and most important characteristic of this interaction is that reinforcement of one person's behavior depends upon the action of another person for its mediation, as contrasted to the situation where the individual interacts directly with the physical environment. The physical environment's reinforcers are more reliable, more immediate. In social reinforcement, reinforcement is unreliable or is based on extremely complex schedules. The basic schedule of social reinforcement is the variable ratio schedule. This schedule of course produces peculiarly high resistance to extinction.

Skinner defines verbal behavior as "behavior which depends for its reinforcement upon the actions of other persons" (Skinner, 1957). Being so broad, this definition subsumes modes of action other than vocal. Linguists might object to such a broad definition but we are all familiar with the fact of nonvocal communication. In

order to treat language, communication, linguistic behavior effectively, we must see it in a broader perspective. The behavior that the teacher seeks to modify then is verbal behavior, according to Skinner. The learning process as it is controlled by programmed instruction differs in no essential way from the learning process as it is controlled in the schoolroom.

A scientist's behavior in studying the behavior of other organisms comprises half of a two-person interaction. It is different from ordinary interactions in that the scientist behaves according to a set of rules that are not immediately modified by the behavior of the subject organism. In the final analysis, the behavior of the scientist is modified by the behavior of the organism *if* his method of analysis is inappropriate. If he is unsuccessful in learning what he seeks to learn, he must modify his approach; he must ask new questions in new ways. In this regard, the behavior of the scientist differs in no way from the behavior of any other person. The behavior of the psychologist is no different in this regard from that of the physicist who studies inanimate matter.

Is not our definition of verbal behavior—or of the two-person interaction which we have taken as our paradigm for automated instruction—so broad as to include all of human behavior? After all, both the scientist and a child exploring some new aspect of his physical environment have their behavior modified by what they do, by the consequences of their actions. This is not a trivial digression, because it forces us to reexamine exactly what element of reinforcement is mediated by another individual. The element peculiar to reinforcement is the scheduling of reinforcement that results from the action of another individual. The physical environment, as has

been noted before, is immediate, reliable, and inevitable. The reinforcement that one derives from the behavior of another organism is no more reliable than the behavior of that organism. The one thing certain about the behavior of organisms is that, to a degree it is uncertain. If behavior were easily predictable, we should not be so challenged as to try to predict it. What is there about the teaching situation that distinguishes it from ordinary social interaction between two individuals? The essential characteristic is that there is an intent on the part of one member to modify the behavior of the other. In this effort the teacher makes use of response differentiation. His objective is the establishment of some specific behavioral repertory in the other individual. The new repertory may only involve the learning of new skills. It may require bringing a complex set of behaviors under the control of subtle properties of the environment. It may involve the establishment of control of behavior by subtle properties of verbal behavior itself.

The differentiation process is highly individualistic. Seldom does one find a report of an experimental study of differentiation anywhere in the literature. Differentiation is what the experimenter does with a laboratory animal before he does an experiment. The experiment usually is concerned with the effects of other variables upon responding. The response is a given. But as anyone who has worked with a rat or a pigeon knows, rats and pigeons do not characteristically go around pressing levers or pecking translucent keys. Before the laboratory animal may be used, the experimenter must condition behavior so that the particular response exists in the behavioral repertory of that organism. So it is not too surprising that the teacher who he seeks to learn from the psycho-

logical literature something about the learning process that he can apply to his teaching, finds little that is relevant. His particular problem essentially is the conditioning of differentiation. We do not normally bother to write laboratory reports that describe in detail how we condition a bird which we then use in, say, a discrimination experiment. This kind of information is the kind of information a teacher needs.

Pacing the Conditioning Process

The experimenter has to pace his own conditioning to the behavior of the learning organism. If we establish too high a criterion in our method of approximation—if we demand too much, too close an approximation to the end product we desire to establish—we may find that we are unsuccessful in conditioning the organism. We proceed by small steps. We must take care *not* to reinforce behavior that is incompatible with the form of behavior we desire. If we demand that the organism move too rapidly we find that our control breaks down. We have to drop back to an earlier stage and begin again to differentiate more gradually. The capacities of the organism are limited. We must recognize that we are imposing a difficult discriminatory task upon that organism. We are giving him a hard problem. His behavior with respect to the problem may be mostly inappropriate. We must recognize, however, that those small beginnings which are appropriate, even though only slightly so, must be carefully strengthened. We must build upon small beginnings. Otherwise we shall not condition our pigeon to peck the key. We will not succeed in teaching a new

concept to a child. As a general principle, it is desirable to move slowly, at least in the early stages of differentiation training. For this reason it has been stated that small steps in programming are desirable (Edwards, 1956). Skinner does not argue that these steps should be as small as possible but insists that the optimal step size is an experimental question. The size of step used must be based upon these considerations: first, it is advisable as a general principle to keep step size small in differentiation; second, differentiation is a highly individualized matter and some organisms can take larger steps; furthermore, for some responses larger steps may be taken because certain responses are more likely to occur in the naive organism than others. The differentiation of stepping with the foot, for example, is a trivial thing for a pigeon. Pulling a chain with the beak, however, is not a trivial matter, and the difficulty of establishing that particular response in any considerable strength is considerably greater. So the paradigm for our learning situation in automated instruction corresponds more closely to differentiation than it does to experiments in the learning process under fixed experimental design.

Modifying the Program

Gilbert emphasizes the importance of working in the early stages of programmed instruction with a single student. He advises the programmer to take the subject matter and to teach it to a student. The programmer then should modify his program, his entire approach to the teaching of the subject matter to that one student. He then teaches it to another student, incorporating the

modifications he derived from his first set of interactions. After this second effort, he has more modifications to make in his teaching process. He incorporates these and uses his process with another student, and another, and another. As Gilbert says, "by the time you have gone through fewer than ten students, you'll have a program that will teach 98% of the students and you will have discovered how to adjust the program for individual differences" (Gilbert, 1960). A good experimenter does this with his pigeons. But he has not written down how to do it, he has not put it in codified form for a machine to carry out.

Programmed instruction differs from the simple differentiation process in that it is a unidirectional affair. The program cannot answer the student's questions. For this reason the construction of the program must be done with much care in order to anticipate many possible questions or problems that may arise in the course of the learning experience. Once the program is in the hands of the student, the programmer can no longer assist it. It must meet all eventualities unaided. It is theoretically possible to build a machine that would correspond to the trainer's behavior to do the differentiation training of laboratory animals. Similarly, such a teaching machine could be built for use with students. It might be instructive to consider for a moment why this hasn't been done. Although an experimenter's time is valuable, such a machine would be even more so. The cost of a machine with the necessary sensing devices to discriminate the behavior of the organism and to react appropriately to it would be prohibitive. A machine could be built for the optimal case where reinforcement is prescheduled independently of the behavior of the organism. Such a ma-

chine would require more than is required of the teaching machine. Although the program presented to the student would be independent of that particular student's behavior, the reinforcement derived from that program would not. Automation of the laboratory procedure is limited by its prohibitive cost. This is a lesson from which we should learn for our later concerns with education.

The Ingredients of Programmed Instruction

What exactly constitutes programmed instruction? First let us define the terms that we use in connection with programmed instruction. The subject matter to be taught is composed into a *program*. The program may be of several physical forms. It may be a book; it may be in the form of tapes or strips of paper; it may be a series of microfilmed slides; it may be auditory material to be used with a tape recorder. It consists of a series of items, referred to as *frames*. A frame is a unit of the program that requires a response of the student. The material in the frame builds cumulatively. Appropriate to the paradigm of differentiation, the program builds in small steps. The information required to answer a given item is contained in that item or in preceding items, or in both. In at least one type of program, the linear program, we attempt to maximize the probability of success. This is done because we conceive of the successful response as reinforcing to the student. One might say that reinforcement, interpreted from the standpoint of contiguity theory, should result from any change the student ac-

complishes. We have already pointed out that this definition presents certain problems as to what constitutes a change, and it might be appropriate at this point to return for a moment to the problem of reinforcement.

Let us propose that reinforcement be interpreted as the result of the organism's reducing a gap between a prior and postaction state of affairs. Learning is a very sensitive and fragile form of this kind of situation. It requires that the organism recognize or discriminate his own state of knowledge and be capable of recognizing a change in it. He must discriminate the adequacy, the range of his own behavioral repertory and discriminate an increment in that range and repertory. Perhaps it is not the recognition that a gap between what is required and what is accomplished has been reduced that is crucial, but that this recognition must be established as a generalized reinforcer in order to be effective. The child must have a sense of the importance of learning. How is this accomplished? Every child has some sense of the importance of learning because he has had the experience that as he matures, as he learns to do more things, his range of control over the physical environment is increased. Irrespective of the theoretical interpretation of reinforcement or of the learning process, success involves extended control over the environment—physical, social, or what have you. This in itself is a necessary, inevitable characteristic of all those things we refer to as reinforcing. If the child learns that learning, of itself, extends this control, then learning will be important. Behavior is important. This is a concept of the first order. It is something that every child learns; it is immediate and direct. Less obvious is that changes in behavior are important

because these changes herald the establishment of even more effective behavioral repertories. Their importance is, however, not immediately evident to the naive. In reality, they are more basically important than simple behavior in that they lead to more effective control. This is a higher order of concept, a more subtle discrimination about behavior than a small child readily makes. This, perhaps, is the crux of the matter for the educator. In making his limited array of otherwise feeble potential reinforcers more effective, the teacher must have some minimal discrimination on the part of his students of the importance to them of learning. This problem is not unique to the first-grade teacher. It is the general problem of antiintellectualism at the college level. At this level, the picture is perhaps more hopeless; the educational process is about to terminate for most students, and there is not much chance that any significant change will be made if an individual has arrived at the age of twenty-one and still has not learned to appreciate the importance to him of intelligent adaptation.

It may very well be that it is in this regard that our system fails as contrasted with the Soviet system. The Soviet system makes clear the importance of a higher stage of learning (Hiebsch *et al.*, 1961). The individual's entire pattern of existence depends on the degree to which he makes use of the opportunities for education. Once this point has been made to the student in terms he understands, it is remarkable how the educational process moves forward. But where success can be achieved more spectacularly in ways that do not involve the acquisition of any particular knowledge or the development of special skills, then that particular knowl-

edge and those skills are not important. Perhaps we should not be disturbed with our school systems for having failed in what is, by the time they are able to reach the student, a practically hopeless task. It is the entire culture that defines what is important. Consider the implications for education of cultural attitudes that attach approbrium to being an "egghead" on the one hand and to being "uncultured" on the other. An amusing sidelight is provided by the statement of an anthropologist that certain nomads in the Sahara were primitive because of the profusion of words in their language having to do with the lion. There was a word for lion, for a sleeping lion, for a hungry lion, for a lion approaching you, and so on. This linguistic state of affairs was taken to indicate that the language of these people betrayed them to be primitive and backward. This is a strange conclusion to be drawn by a member of a civilization whose own language enumerates various kinds of automobiles in even greater profusion. The fact is that the lion is an important feature of the Bedouin's environment in the Sahara and, as such, he makes important discriminations about it that must be packed into a small number of words in a very small period of time for the sake of communication. We have other discriminations about other things in our environment which are of central importance. There is nothing primitive about this feature of language in either case. If we seek to determine what is important to our culture, let us look to those features of our environment for which we make the most refined discriminations. Education and its accompaniments do not usually fall into this category.

The Principal Features of the Teaching Machine

There are three important features of the teaching machine. First, it presents material to the student in an organized, logical sequence. Second, it requires from the student an overt response. Third, it provides feedback to the student so that he knows whether or not his response was appropriate. Any device or system that is composed of these three elements is a teaching machine. What about books? Books present material in organized, logical sequences and students do sometimes make overt responses to them. The customary overt response to a book is to sit down, open it, orient the eyes toward the page, and turn the pages. Covertly we read. There is no direct indication that reading occurs during the time of its occurrence. We are all familiar with having sat down to read a book and, on page 83, suddenly realizing that we don't know how we got there. We turn back 2, 3, 4, 5, perhaps 6 pages before we recognize material which we had previously read. We have been mechanically going through the process of scanning material and turning pages, but that material is not in successful control of our own verbal behavior. The requirement that the student emit an overt response insures that this form of woolgathering does not occur. If it occurs, the student is immediately aware of it. It produces a sharp, clear break in the studying process. To be sure, there is implicit reinforcement in reading a book, but if the reinforcement to be derived from the more direct and immediate class-

room situation is ephemeral, then those subtle reinforcers implicit in having "learned something" from a book must be even more so. The lecture suffers from the same limitations, with the additional one that the lecturer may present his material in a way that is more conducive to hypnotic slumber than to the establishment of the effective control over the behavior of his students that the lecturer intends. The same can be said of most audio-visual aids, such as movies, filmstrips, and so forth. No overt response is required, and there is no immediate monitoring of behavior. Those with experience in military service will recall that the primary task in watching training films was to stay awake.

Many persons view the place of programmed instruction as similar to that of other audio-visual devices, but the characteristics of these various techniques differ vastly.

The teaching machine is not simply another audio-visual aid. It represents the first practical application of laboratory techniques to education. The task of programmed instruction—as of all instruction—involves the conditioning of a behavioral repertory. We do not conceive this activity to be the instilling of a body of knowledge, although what we do has an effect that is described by that phrase. We seek to increase the behavioral repertory of the student. We seek to establish a complex class of behaviors and to bring that class of behaviors under the control of particular features of the environment. This is the task of teaching, whether it be directed to the rat, to the high school sophomore, or to the graduate student in philosophy.

The Application of Mathematical Concepts

We have occasionally made reference to mathematical learning theory and to methods of analysis taken from mathematical learning theory. The teacher can make as profitable use of certain of these concepts as can the experimental psychologist who studies learning in the laboratory. To speak of stimulus elements is to conceptualize the environment, to use a kind of shorthand. We need not assume that there are really such things as stimulus elements. But if we do no violence to the facts, and if it does give us an increased ease in communication to speak of such elements and to use them in our description of the teaching situation, then their use is legitimate. If other conceptualizations are more effective, the present ones should be abandoned in their favor.

One concept that comes from mathematical learning theory is of importance because it is relevant to questions we shall consider with reference to the function of different types of programs. This concept is the so-called equal-alpha assumption. This says that the extinction of response $\bar{A}$ is equivalent to the reinforcement of response A. In experiments testing propositions of statistical learning theory, a particular experimental setting is employed which forces the subject of the experiment to make one of two mutually exclusive responses. This is done to bring the realities of the conditioning situation into correspondence with the assumptions of the theory. It will be recalled that conditioning and extinction in terms of mathematical contiguity theory are regarded as

transfers of response probability from one class of behavior to another. To force the individual into a situation where he must make one of two mutually exclusive responses is to limit his behavior to correspond to the conditions about which theory makes predictions. This simplification is a legitimate experimental procedure, but the implications of the assumptions involved should be examined as we leave that rather confined circumstance to view a more general situation where the behavior of an individual is not constrained in this way. What happens to behavior that is extinguished in the conventional teaching situation? Theory says that the extinction of $\bar{A}$ is equivalent to the reinforcement of A. The increment in response probability resulting from this operation is equivalent to the increment in response probability of reinforcement of A. In certain forms of instruction the teacher tells the student that he is wrong. *Some* form of feedback occurs. Or information may be given the student through lack of feedback, as in extinction. In a sense, the lack of feedback in extinction amounts to the same thing. Either the subject is reinforced or he is not. Both extinction and being told that he is wrong let the student know that his particular response was inappropriate. What is the effect upon the *appropriate* response of this kind of teaching procedure? If the behavior of an individual were limited to two mutually exclusive response classes, telling the student that he is wrong might have an educational effect comparable to telling him that he is right when he is. We know that telling him he is wrong is aversive to a student. It has effects that compete with learning. Being told that one is wrong in a competitive social context, as in a classroom, must produce an emotional response in any intelligent child. Re-

spondent behavior is conditioned to the classroom, it is conditioned to the teacher, and that respondent behavior will interfere as long as it lasts with more adaptive behaviors. Aside from having produced maladaptive emotional behavior by such a practice, we also have failed to lead the student to do the right thing. His behavior is not neatly organized into two mutually exclusive classes. When a child makes a wrong response to a problem situation and is told, "No, that's not the thing to do—do something else," he may resort to any one of a large number of alternatives with respect to that problem, only one of which presumably is correct. Although in theory telling a student he is wrong may, in a way, be opposite to reinforcement, in practice this is not the case. Telling him he is wrong may produce an increased likelihood of the correct response; it may produce a decreased likelihood of the particular incorrect response, but the increment in probability is spread over the entire range of alternatives. Not only is there an increment given by this procedure to the correct response class, but there is also an increment given to all the other possible incorrect responses. The probabilities are increased that the child will do something else, but they are not changed in favor of his now doing the correct thing. This consideration is going to be important to us when we examine certain methods of instruction in the next chapter.

To summarize, a program consists of a series of stimuli designed to exert increasing control over a gradually developing behavioral repertory. Reinforcement derived from matching behavior to the stimulus requirements accomplishes the differentiation procedure. The frames of the program are discriminative stimuli anticipating probable courses of action on the part of the stu-

dent and directing his actions by limiting the range of alternatives available to him. They are instructions in the broadest sense. Violating the limits results in nonreinforcement. Remaining within the bounds results in reinforcement. The bounds narrow progressively to define the ultimate form of the new repertory. As they have been developed, programs make use of continuous reinforcement, both in the development of discrimination with regard to the material to be taught and in the differentiation of the response class to be shaped. In this respect, they embody the means for very rapid training. They also open possibilities for applying other techniques of control to produce more permanent learning.

CHAPTER SEVEN

Teaching Machines

In a sense the techniques of programmed instruction have been in existence as long as learning has existed.[3] One can trace the development of the first device labeled a teaching machine to Pressey in his paper published in 1926, but patents were issued by the U.S. Patent Office as early as 1809 for devices aimed at aiding teaching (Mellan, 1936).

Pressey developed a series of devices and has for many years continued research in the application of automated teaching techniques. The influence of his work has been strongest in industry and in the military services. The military services have given a great deal of attention to the development of instructional devices derived from Pressey's techniques. Flight simulators,

[3] The cover illustration of the *American Behavioral Scientist*, Vol. IV, No. 9, shows what is called "A Medieval Teaching Machine." This is the quintain, which was used for the training of knights. The response appropriate to this device was striking a shield directly in the center with a lance. If the blow was correct, the device fell over. If the blow was struck off center, the device would pivot and deliver feedback by striking the horseman a blow with a flail or some other instrument as he rode by.

trouble-shooting trainers, tab items, and other devices are representative of these efforts. The technique of intrinsic programming was developed for early work that Crowder did while employed at the AFP TRC Laboratory. Pask, who also views the teaching machine as one side of a two-person interaction, came to the area of programmed instruction through the route of automatic control, cybernetics, and self-organizing systems. Skinner's contributions, although continuous with these earlier efforts, represent a real departure in technique and concept. The earlier developments are organized around the concept of individual differences. Skinner's work has grown out of the concept of conditioning.

The Ideal Machine

We noted in the preceding chapter that there are three essential characteristics of the teaching machine: the sequential presentation of material, the requirement that the student make an overt response to the material, and an immediate feedback to the student informing him of the adequacy of his response. What is the physical nature of a device that accomplishes these things? If we follow our paradigm for programmed instruction as it is represented by two-person interaction, then the ideal teaching machine is a tutor. The tutor would be a machine of great complexity, a machine that could take into consideration the peculiarities of the individual student and modify its own behavior, its own functioning, in order to work most effectively with that particular student. The tutor would be an experimenter who had enough

time to differentiate the behavior of all of his subject organisms, even where that behavior was composed of a large body of information and where the time required for its assimilation might consume years. This would be an expensive device. An approximation of this complex device is the picture of the ideal teacher that Galanter gives us (Galanter, 1959). Galanter proposes a mechanism that would tax the capabilities of some of our most advanced computers. The machine he proposes is possible; with certain modifications it exists today. It is as expensive, or perhaps more expensive, than a flesh-and-blood tutor would be. Both of these examples make the point that one of the limitations of a teaching machine must inevitably be its cost. The economics of programmed instruction provide us with one of the most interesting facets of the entire business, both from the standpoint of the feasibility and from the standpoint of the probable course of development of the field. In a way, what Galanter proposes is a Turing Machine.[4] The Turing Machine is a device whose operation is indiscriminable from the operation of a human. We are not, as Turing was, concerned with whether or not the machine thinks. We are concerned with the capabilities of its responding to the behavior of the student as a human teacher might respond. Is the machine capable of modifying itself to correct errors, to change the pace, to alter the subject matter in response to the needs and abilities of the learner? The human is capable of this. A machine can be built that also is capable of this. Clearly, the little

[4] A. M. Turing, "Can A Machine Think?" *The World of Mathematics,* Vol. IV, p. 2099. Paper originally taken from *Mind,* 1950.

red schoolhouse in the small rural community will not be in a position to purchase a machine with such capabilities. What compromise, then, can we make with costs and still achieve a useful device? We must sacrifice flexibility of the device in the interest of the budget. Just how much flexibility can we sacrifice and still have an effective teaching device? This is a question that must be answered in the context of the resources of a specific school system. It is unfortunately true that some school systems will have to compromise more than others in this regard.

The Machine as It Exists

So far we have discussed theoretical teaching machines —machines that do not exist. What can we say about the complexities and capabilities of those machines which have been constructed? We shall not undertake an exhaustive survey of the hundreds of devices because such a review is published elsewhere (Foltz, 1961) but we shall touch descriptively upon a few examples varying in cost and complexity. In terms of capabilities, perhaps the most elaborate machine is a device that can present material visually in the form of panels—that is, pages of material, items, frames, filmstrips, moving pictures and sound. The only modalities that have not been programmed into this device, at least at this point, are smell, taste and touch. The cost of this machine per unit is high. It is so high that purchases, when made, are usually made by governmental or industrial concerns. Even then procurement is usually on a rental basis.

In descending order of complexity, we might consider the first model of the write-in machine that Skinner used. This machine presented frames through an opening on the top of the device. The program was originally printed on a series of disks. This is the so-called disk machine. It had a drop-out device that permitted the repetition of those items of the program that the student missed. This is the machine with which the memory drum was compared in Chapter Five. The drop-out device is an expensive feature in the write-in machine. A less expensive machine has all the features of the disk machine with the exception of the drop-out device. Such a machine presents material in a straight-line program. It also permits the programming of supplemental prompts. If the student is unable to answer a particular frame on the basis of the information in the frame he may reveal a prompt to himself, that is, reveal additional information that will facilitate his correct response. After he makes his response he closes the response space so that he cannot change his answer; he is then shown the correct answer. An additional feature of this machine is the fact that it locks and cannot be opened by the student. This particular machine has a self-scoring device that tells the student how many items he missed and also marks the program tape in such a way that the programmer readily gets an indication of the items that cause students the greatest difficulty.

The next machine in decreasing order of complexity and cost employs a straight-line program but has no prompting feature and no cumulative score counter. After he has responded to a particular frame, the student is able to score himself by advancing the tape, moving

the frame and his answer tape so that he cannot change his answer. He also thereby reveals the appropriate answer to the frame for comparison. Still lower in degree of complexity are the various forms of punch boards which are to be used in conjunction with multiple-choice programs. The student is given a punch board that is keyed to the particular set of items comprising the program to which he must respond. His response consists of punching out holes which he believes represent the correct alternatives to the particular items. He continues to punch until he punches the item by which he is informed by whatever means is the correct one. The material may be programmed sequentially. He must make an overt response and he does get feedback. There are theoretical objections to the multiple-choice kind of program, but insofar as satisfying the basic requirements of teaching machines is concerned, the punch board does qualify as a teaching machine.

Several specialized devices for the training of simple visual discriminations, the training of matching to sample, are described in other works, but because we are concerned with programs that are somewhat removed from the establishment of simple visual discrimination, we will not consider such devices in detail. These devices are especially useful with preschool children or children with particular sensory or other deficits. They are effective in dealing with preverbal or nonverbal subjects. Related devices are used in the training of specific skills and dexterity, the operation of keyboards, etc. These are mechanical devices that respond to the learner's rate of responding and to his frequency of error and adjust the requirements of the task in accordance to these (Pask, 1958).

The Programmed Textbook

Finally, in the realm of teaching machines we should consider the programmed textbook. The programmed textbook is a program that does not utilize an auxiliary device for its presentation. In all cases, it must be emphasized that the machine per se does not teach. The machine is simply a device for presenting materials. The student learns these materials through his interaction with them. The machine is an adjunct to this process. Some would argue that it is not a necessary adjunct, that it may, in fact, be detrimental. Many persons in the field favor the programmed textbook. As we know, the programmed textbook presents a frame, the answer to which appears on the next page. This is the programmed textbook as it is used with linear programs. An alternative to the linear programmed textbook is the scrambled textbook based upon branching programs developed by Crowder and his associates. In these, the student is given an item and a set of alternative answers. Depending upon the answer he chooses, he is directed to another page for the next step. If his answer was correct, he moves on through the program. If it was wrong, the material on the page to which he was directed informs him of this and explains why he was wrong. It then directs him back to the item he missed for another try. Both linear and branching programs have been successfully published in textbooks of these types.

The programmed textbook, in a way, has been around for a long time in the form of workbooks. Granted, workbooks are not usually carefully worked out

concepts; sequential programming does not exist. Overt response to an ordered series of items is required by the workbook and it characteristically has a set of answers with which the student can check his own. One of the greatest drawbacks to the workbook is that the student can glance over the entire presentation. From the context of succeeding items he can fairly easily figure out the answer to a particular question. Any advantage that may be derived from the gradual, orderly differentiation of a response class is lacking. The arrangement of material in the programmed textbook overcomes this handicap to a great extent, but it is still possible for the student to move back and forth. If he comes to an item he cannot answer, it is a simple matter to flip ahead to the next page for a quick peek.

To be sure, a student who knows that this kind of cheating defeats the entire purpose of programmed instruction will be less likely to try it. After all, the only person he cheats is himself. Yet, after he has worked for a long time in such a textbook, he becomes fatigued. The temptation to cut a corner here and there does exist. Moreover, as the student idly scans a page of frames in a programmed textbook, he often spots an item ahead of his place that he can answer. The tendency to answer it is great. The feeling that he is getting ahead of himself provides reinforcement.

The teaching machine potentially gives more adequate control over the sequence of the learning process than does the programmed textbook. It may well be that this additional control is not particularly important, but until more evidence is available, we shall not assume that it is unimportant.

Criticisms of the Machine

Where a particular program has been demonstrated to be effective in teaching a subject matter, the teacher may choose to base his grading system upon the completion of the program. For this reason, some of the write-in machines have locking devices so that the student cannot tamper with his answer tapes. Once the grading of the student is made contingent upon his response to the tapes, policing of instruction becomes important; locking the machine becomes important. If a teacher anticipates the use of programmed instruction in such a context, the programmed textbook, of course, is inadequate. At the present stage of refinement, many telling criticisms can be and have been made of machines. Existing machines are unreliable. They break down. They are not standardized. A program that fits one does not fit another. What is more important, the existing machines may dictate the kind of program to be employed (Gilbert, 1960). Such instances of the tail wagging the dog are a major roadblock to the progress of this entire effort.

A different kind of objection to machines is that their effectiveness depends upon their novelty. It is agreed that at first the student is motivated to work with the machine because he is fascinated by the gadgetry; he regards it as a kind of toy. It is argued that the amusement thereby provided constitutes the real reinforcement for working with it. If true, this is a telling criticism, for we can expect that to the sophisticated student of 1984 who has received all his primary and secondary education on machines these devices will no longer be novel. It is unlikely that the effectiveness of the machines de-

pends exclusively on this aspect, though it must be granted that fascination with their novelty is a significant component of the reaction of naive students facing these devices for the first time.

Realistic Uses of the Machine

What is the realistic concept of the teaching machine in education? What are reasonable objectives for its use? Should we seek to program all knowledge? Can everything be taught in this way? In theory, the answer to these questions is yes. It is possible to program anything. In practice, the answer is no. A machine that could teach laboratory analysis in chemistry, for example, would be of prohibitive complexity and cost. Further, the time required to build a program to teach material of a very abstract or subjective nature is uneconomical compared to the time required to teach it in a standard way. It is not the objective of any program to replace the teacher. The book did not replace the lecturer. The program will supplement the teacher. The program does not make the task of the teacher easier, for if the program is used effectively, it will provide the teacher with a classroom full of students who are as familiar with the basic didactic material of his subject as he is. In order to justify his continued existence in the classroom, the teacher is required to add something to the basic facts. It is his task to stimulate, to bring imagination into the classroom, to integrate, and to speculate. It is the teacher's task to show the students how to manipulate the basic facts, the items of information that constitute his subject matter so that

they learn new things. It is the teacher's task to teach enthusiasm by example.

Should the program examine the student? This depends upon the objectives of the particular teacher and on the particular course. The machine can examine, but the program itself should never be conceived of as a testing device. It is a teaching device. One of the greatest pitfalls for the inexperienced programmer is the tendency to construct examinations rather than programs. Machines teach, although they may be used as testing or review devices. But to relegate programmed instruction to either of the latter two uses is to miss the point completely. The machine is the teacher par excellence for certain material. It is a misuse and a waste of both machine and teacher if either usurps the other's particular domain of greatest competence. The machine is not just another teaching aid. Some commercially marketed devices link the individual responses of students taking the course to a common presentation of visual or auditory material. As yet, there is no way to allow self-pacing with such a teaching system. The educator should be aware of this limitation of the system. Such devices are relatively more expensive and, though technologically impressive, have severe limitations from the point of view of individual flexibility. With such a system, the entire group is paced to the rate of the slowest student in the class.

Should the machine be used to pace the student? One of the essential features differentiating the teaching machine from the memory drum is the nonpaced nature of the response to the teaching machine. The student proceeds at his own pace. This is an advantage for the individual student. It may be a mixed blessing for the

teacher. What does it do to the classroom? What happens in a supposedly homogeneous group when the brightest student completes the material in a third or less of the time required by the slower student? What implications does this have for the design of the curriculum, for the grading of students? One of the most far-reaching implications of machines has to do with the burden the technique places upon the teacher and the school administrator to exploit the technique to its fullest. The very nature of the operation of programmed instruction must inevitably produce such ranges of divergence in skill as to create acute administrative problems within our present educational framework. These problems should be welcomed rather than avoided; in their solution lie the answers to more serious problems that challenge our present way of life.

CHAPTER EIGHT

Techniques of Programming

The first consideration that determines the form a program is to take is the nature of the material to be programmed. Programming is writing a book. Appropriately enough, programs have been constructed to train programmers to program (Mechner, 1961; Markle *et al.*, 1961). Rigney and Fry have recently published a report where the technique of programming is presented by example (Rigney, Fry, 1961). These materials fulfill a useful function, but in the final analysis, the training of the programmer is done by the student. No one can tell another person how to write a good book. No one can tell the teacher exactly how to write his program. The only way to learn to write a specific program is to write it. The would-be programmer who seeks a set of rules telling him how to write his particular program will be disappointed by the present analysis. In all honesty, no one can prescribe the rules for the successful writing of a program in a specific area. The final judge as to whether or not a program is effective is the student.

What we can do is to set down those general prin-

ciples of programming which are known to be important at the present time. We can prescribe certain procedures and techniques of evaluation that will tell the prospective programmer how to find out whether what he has written is a good program. We can examine techniques of programming to see what have been features of good programs and we can suggest ways by which programming may be simplified. Writing a program is behavior. As behavior, it undergoes differentiation as a result of success or failure. The programmer learns to program. That which modifies the programmer's behavior is the consequence of his efforts.

The prospective programmer should not assume that he knows how the student learns. The programmer should approach the student with the same degree of open-mindedness with which an experimenter must approach a laboratory animal. The programmer initially knows nothing about the potentials of the particular organism with which he is about to work. The organism will show him what his potentials are. The organism will, by his success in learning what the programmer desires him to learn, unequivocally show the programmer how successful the programmer is in arranging materials to be learned.

The Linear Program and Branching Program

At present there are two basic types of programs. The first, and the one with which we shall be most concerned, is the straight line or linear program as developed and used by Skinner and his associates. It is the program best

suited for use on the simpler teaching machines now in existence. A linear program is composed of small steps leading logically through the subject matter from topic to topic. This type of program is based on the kind of behavioral analysis that we have developed. In this type of program it is important that the student make as few errors as possible. To this end the increments in information which the student is expected to absorb are small.

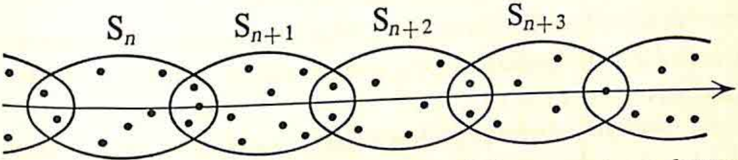

FIGURE 7. Schematic representation of linear program frames regarded as a series of overlapping discriminative stimuli distributed in time. Dots within frames indicate elements composing stimulus subsets.

A series of frames may be viewed as a sequence of stimuli or stimulus elements sharing some elements from frame to frame. One may regard learning as the conditioning of behaviors to the elements within a frame. Through reinforcement, the probability of a correct response is increased to those elements within a specific frame; one then moves to the next frame. The probability of response to the next frame is higher than it otherwise would be because some intercept elements are shared with the previous frame or frames, to which the response has already been conditioned. Thus, the organism proceeds from the known to the unknown. The sequencing of elements in this manner is exemplified by the diagram in Figure 7. Stimulus discrimination is schematized by a pair of adjacent sets of elements. S^D and S^Δ share certain intercept elements. The continuing process of dif-

ferentiation is indicated by chains of discriminative stimuli overlapping with one another in terms of the elements shared between successive frames or discriminative stimuli. Clearly, it is redundant to have too large an overlap between successive frames. At the same time, a certain minimal overlap must exist in order that the probability of response to succeeding frames be high enough. Frames in this sequence must not be overly dependent upon one another nor may they be completely independent of one another. These extremes, of course, exemplify programs that are too simple and programs that are too difficult.

Certain objections have been raised to the Skinnerian position that step size from frame to frame be small to increase the probability of reinforcement. Markle has shown learning to be more efficient if errors are reduced or eliminated. Let us consider the arguments against the notion that learning is more efficient if errors are reduced or prevented. The entire philosophy of small-step programming is based upon this assumption. One detects in the objections to this position the notion that no one has demonstrated that constructing a correct answer to an item is reinforcing. Operationally, there is no question but that this is reinforcing to the student if he learns thereby. The problem exists in giving reinforcement hedonistic connotations. We have already dealt with this problem in Chapter Three. Proponents of linear programming do not imply anything about successful responding beyond the fact that it strengthens behavior. Nothing is said concerning drive reduction. The confusion arises from a misunderstanding of reinforcement. If reinforcement is used as the layman uses the word re-

ward, then perhaps a problem exists. If the term is used operationally, it means simply that a student's answering an item correctly results in learning. This is reinforcement of exactly the same essential nature as that resulting from giving a rat a pellet of food for pressing a lever. It is a strengthening of behavior. This is a definition with which no one can argue, irrespective of his own theoretical predilictions. If someone becomes involved in surplus meanings and theoretical explanations, that's his problem. Such theoretical implications are not a part of the concept or of the technique as it is employed in the linear program.

The second basic type of program is the so-called branching or intrinsic program developed by Crowder (Crowder, 1959). In the branching program, the student is presented with a problem and with several alternative answers, one of which is correct. When the student chooses an answer he is instructed to move to a specified frame. This frame then tells him if his answer was incorrect and explains why it was incorrect. The frame then may return the student to the original item which he had answered incorrectly for another trial, or it may direct him through a subprogram—further instructing him in the basic knowledge presumed to be necessary for the item he had answered incorrectly. In either case, the student is eventually returned to the missed item which he then, presumably, answers correctly. If he again chooses an incorrect alternative, a similar process is followed. Ultimately, he is returned to the missed item and answers it correctly. He is then directed to the next frame in the program where the same process may be repeated should he answer that item incorrectly. Both the

linear program and the branching program lend themselves to publication in the form of books. Several books have been published using the linear format. Branching programs are published in the form of scrambled textbooks. The scrambled textbook arranges the frames and alternative answers to frames in such a way that the student is directed to search through the book to proceed to the next step. It does not move sequentially, page by page, as does the linear program.

Analytically, the branching program is a first cousin to the multiple-choice examination, in contrast to the linear programming technique where a response must be constructed. The student is faced with a recognition task in the branching program where he chooses from a set of alternatives presented to him. The range of possible responses is mechanically limited in the branching program. Alternatives must be limited in this format as it is impractical to anticipate all the possible responses a student might conceivably make to a particular item and to write replies to each. A program that incorporated all possible answers to each step would exceed all reasonable limits of size. The branching program departs from the conventional multiple-choice examination both in the immediate feedback feature of the program and in the explanation of why a particular item is incorrect. One well-known objection to the multiple-choice examination has been that the student may learn (if, indeed, he learns at all from examination) an incorrect answer to a question. Presenting three times as much erroneous material as correct material to the student raises the likelihood that this erroneous material will ultimately be retained instead of the correct associations that the student should establish. Similar objections may be raised

to the branching program, though these objections are vitiated by the immediate feedback in the form of explanation incorporated in the program.

Crowder and others argued against the position that learning of programmed material follows principles that apply to the laboratory conditioning of experimental animals. Instead, he argues that the true paradigm is the student-tutor relationship where the tutor leads the student to discovery, employing the Socratic technique with immediate feedback to the student's proposed answers to problems. Unquestionably, the interaction between programmer and student as it exists in the impersonal format of the program is of this type. However, unless something is implied concerning an emergent process generated by the interaction of two people, it is difficult to see what interaction—even between Socrates and Plato—could be, over and above the interactional processes we have discussed. We, too, take the position that the teaching situation consists of this interactional setting. We have arrived at this position from learning principles. The only conclusion we can draw is that the objections are addressed to the indignity some persons may feel they have been subjected to when it is suggested that they share common behavioral processes with the "lower" animal.

Implicit in Crowder's position is the notion that the linear program is an insult to the intelligence of the brighter student. By moving the student along in small steps, we somehow do not call forth or demand of him his fullest potential—according to Crowder. The reasoning required of a student in a branching program may be considerable. This may also be true of the linear program. Simple, slow-moving programs of any type are an insult

to the intelligence of a bright student. If not an insult, they may, at best, be relatively unstimulating to the brighter student. Programs are like ordinary books in this regard. We cannot agree, however, that Crowder's technique of implicit programming involves essentially different processes of conditioning. They are simply more contaminated by the complexities of interaction. The same processes of conditioning must inevitably be involved in both cases. Implicit programming may set certain limits to the range of behaviors that are possible and it may create certain problems in the analysis of the performance of students, but its objective is the establishment of a verbal repertory. It is essentially identical to other techniques of verbal conditioning. There is no discontinuity here.

We agree in one sense with Crowder's insistence that the student be given a chance to be wrong. If a frame so constrains the range of possible responses that the correct one is inevitable, then probably the student doesn't learn much from making such a response. After all, there isn't much reinforcement, generalized or otherwise, to be derived from busy work. If generalized reinforcement is derived from making a correct response, if it is derived from having accomplished *something*, then the *something* must be worth while. The trivial correct response cannot be expected to contribute much by way of reinforcement. We cannot agree, however, that the sheer existence of error necessarily accomplishes any educational objectives. Moreover, the branching program provides less opportunity for meaningful error in any useful sense than does the linear program. The development of an effective linear program is a far more challenging task to the programmer than the develop-

ment of an effective branching program. The branching program creates a small set of straw men that are easy for the programmer to demolish because he tends to select those straw men which are demolishable. The linear programmer, however, must anticipate a greater variety of such straw men. He is not allowed to attack these directly because of the very structure of his technique; he must instead guide the student to the desired response. For those who argue that the principle objective of education is teaching the student to discover new things for himself, the linear program is by far the more satisfactory instrument. It is also by far the more difficult instrument to construct.

One final comment is in order with respect to the kind of response demanded of the student. Let us compare the behavior required of the student in the linear and branching programs. It is well known from early studies on verbal learning that the poorest performance in a test will be obtained from the so-called *construction* method. This is the method whereby a student is required to reconstruct the material without external stimulus support. The method of testing that gives the highest scores is the so-called *recognition* method. It is of interest to note that the linear and branching programs involve these two extreme response modes. The linear program requires an act of reconstruction. The branching program requires recognition. But if a student has learned a set of material so that he is capable of constructing that material, he has certainly learned it well enough to recognize it. The reverse might not be true. In addition, the interpretation of error rate between the two procedures may differ. The meaning of an error rate of 10 percent in constructed items is different from the

meaning of an error rate of 10 percent where the subject is presented with a set of items composed of 10 alternatives each among which he may choose at random. The point is somewhat exaggerated because no branching program offers this large a number of alternatives. But to the extent that random sampling does give a lower error rate for the branching program than is true of the linear program so should the criteria of difficulty be interpreted and adjusted accordingly. The probability of making a wrong response where the correct alternative is 1 of perhaps 100 reasonable responses is much larger than the probability of an incorrect response where the correct answer is 1 out of 3 alternatives. These objections are raised simply to call attention to the nature of the task required of the student and to the implications for the interpretation of results drawn from the two techniques.

First Steps in Programming

What must be done first in programming a course, whether it be done as a linear or as a branching program? The first task is to determine the scope of the subject matter to be taught. How much detail does the programmer wish to include in the program? We must consider the gross amount of material to be covered, the time within which that material may be covered, the length of the program necessary to cover it, and the proposed relationship of the program to the teaching environment in which it is to be employed. Is the program to be used to the exclusion of all other techniques? Seldom will this be true. The objective of programming is

not to replace teachers but to provide a basis from which the teacher can proceed.

Having determined the level of completeness he is aiming for, the programmer must translate the concepts to be learned into behavioral terms. What does the programmer wish the student to be able to do upon completion of the program? What verbal skills should the student have? What questions should the student be able to answer? What discriminations should the student be able to make? Having made a list of these objectives, the programmer has defined the composition of his program. He then must determine if there is any logical sequence that suggests itself as the most efficient way of organizing the program.

Once he has constructed such an outline, the programmer will be faced with a series of terms to be defined and with the stating of their relationships. A technique called the *ruleg system,* designed to aid the programmer in covering all the contingencies of his subject matter, has been proposed by Evans, Homme and Glaser (Evans, Homme and Glaser, 1960). They suggest listing the concepts, or, as they refer to them—rules which are essential to the program—on intersecting horizontal and vertical axes. This arrangement they call a *ru* matrix. Once the matrix has been constructed, the programmer is to collect examples of concepts in all their possible interrelationships with one another as revealed by the matrix. Once these are ordered in a logical sequence, the programmer may proceed to write examples that define his concepts. His examples will employ the terms he wishes the students to learn, both in a simple expository form and in relationship to one another. The complete definition of a concept depends upon the program-

mer's supplying many related examples. It has been said that programmed instruction tends to concretize the learning of concepts. If a program is too narrow in its range of definitions, there is a risk that the student will have difficulty in generalizing from the particular items to which he has been exposed to new situations to which those items of information are relevant. This is not an exclusive characteristic of programs. This problem exists for all forms of teaching. It is explicitly evident in programs, and this very explicitness is again an advantage. The student can be taught to generalize by the skilled marshalling of examples.

As a programmer gains experience in programming he may find that the student's response requires the insertion of frames that force discrimination between concepts. It can be done by inserting a sequence of contrasting examples. This is often necessary where discrimination is based upon very subtle properties, as in the definition of abstract concepts. This procedure amounts to a series of alternations between S^D and S^Δ occasioners in discrimination training. As will be recalled from our discussion of discrimination learning, there are important considerations of the number of presentations and proportions of S^D occasioners in the series. The experimental literature here offers a great deal for the continued development of the technique of programming. A converse procedure may be followed to develop the generality of abstract concepts. A series of frames is developed that draws together examples of the concept that differ in various particulars but which define the concept in terms of the property common to all the examples.

This brings us to an important point concerning the background of the programmer. The person most able to

develop an effective program in a given subject is the person most familiar with that subject. It has been truly said that it is easier to teach a physicist programming than it is to teach a programmer physics. For this reason, the person responsible for teaching the material is the person who should do the programming.

Supplementary Devices

Where a program is to develop a verbal repertory, the programmer should never lose sight of the fact that the structure of language is itself an invaluable tool in the development of programs. Supplementary devices may be used in conjunction with a program, whether linear or branching. The commonest of these is called a *panel*. The panel usually consists of a graph or table, or other supplementary information that enables the student to deal effectively with the frames of a particular portion of the program. The panel as such constitutes what has been called a *prompt*. A prompt is a device that increases the probability that out of a set of alternatives a particular response will come to be emitted. It is a pure discriminative stimulus. It depends upon the prior establishment of stimulus control for its effectiveness. Skinner has categorized prompts into formal and thematic prompts. The prompt differs from what is called a *probe* in that the prompt is employed by an experimenter where he knows or can identify the response he seeks to evoke. The experimenter may not be able to identify the response he seeks to evoke with a probe. Both the prompt and the probe depend upon a prior history of conditioning which the other person in the interactional setting

can reasonably assume. *Formal* prompts may involve either echoic or textual prompting where the stimulus is of the same form as the response to be evoked, as in prompting at the rehearsal of a play. Reading textual material is another formal prompt very like the echoic prompt. Abbreviations constitute yet another class of formal prompt. The *thematic* prompt is better known as a hint. The thematic prompt differs from the formal prompt in the directness with which it is employed. The formal prompt is a direct reminder whereas the thematic prompt may be concealed in other verbal behavior.

Language is not a random process. A chain of words in a phrase or sentence determines to a great degree the next word to follow. The determination may be evident only in the form or class of word that is determined. A sentence such as "We are going to the ____" determines that the word which completes the sentence must be a noun. The sentence "We are going to ____" allows that the missing word might be either a noun or a verb. Limiting the choice of words to either nouns or verbs considerably reduces the range of possible responses with which the sentence can be completed. In a real sense, the entire interverbal dependency of language may be regarded as a kind of thematic prompt. In fact, the linguistic structure may so strongly determine the response that the programmer must occasionally take steps to broaden the range; the example "The stimulus evokes a ____" not only narrows the range of verbal alternatives to nouns but also narrows it to that class of nouns which begin with consonants. If it is felt that this additional restriction makes the answer trivially simple, the programmer might wish to write the item "The stimulus evokes a (n) ____." Here the response still must be a noun but with-

out the further restriction that the noun begin with a consonant. The thematic prompt is a specialized form of interverbal dependency. Since the structure of language inescapably provides some prompts that determine a response, the programmer must be wary of using excessive prompts. Excessive use of prompts makes a response trivial. Although the error rate should be held to a minimum, this should be accomplished without sacrificing the necessity for the student to think. An item should not be made so simple that a student can respond automatically without any effort. In this we are in complete agreement with Mechner's view that a low error rate is but one feature of an effective program (Mechner, 1961).

Another example of the exploitation of interverbal dependencies is in the construction of items containing parallel grammatical constructions. Strong tendencies exist for humans to emit synonyms or antonyms to verbal stimuli. For example, in studies of free association, many common responses to verbal stimuli are the opposite or the antonym of the verbal stimulus. One may exploit this tendency in the verbal behavior of students. Consider the example "With increased food deprivation the response rate of the organism increases and with decreased food deprivation the response rate ——." Not only does the tendency to emit opposites determine the response in this item, but the parallel construction strengthens the tendency to echo the word "decrease." In the first half of the statement "increase" is repeated; the parallel construction determines the repetition of "decrease." The textual prompt is probably extensively used in programming for several reasons. A new term must be introduced into the program context before the student is required to emit the term as a response. Some repetition is in-

evitable. If the term happens to be a basic concept whose definition is one of the objectives of the program, it must occur in multiple contexts as various examples are brought to bear in its definition. The repetition of key terms is desirable from the standpoint of learning theory. The particular response should be evoked under as broad a spectrum of conditions relevant to that response as possible. This is another way of saying that a program should be general as opposed to overly specific.

Two comments relative to the effectiveness of various prompts are in order. First, students tend not to make use of panels. Students will use panels only where the text of the program forces them to it directly and repeatedly throughout the relevant portion of the program. Second, students tend to ignore information following a blank in a frame. In items where there is important information relative to subsequent items (this includes almost all items), the programmer must guard against placing any significant amount of that information after the last blank in the frame. It seems that students ignore this information for the same reason that they pay little attention to panels. The student reads the item or panel until he gets enough information to make the immediate response, then he moves on. If information that is important for subsequent items follows his response, he is very likely to miss it in his eagerness to proceed to the next item (Meyer, 1960). Markle specifically excluded the form of prompting she referred to as the "hoary educational technique of 'repeat after me.' "

A related technique has been employed experimentally with certain types of materials, but as yet it is too early to say whether this technique of programming is more or less effective than more conventional

techniques. Holland has proposed that a technique known as "fading" or "vanishing" would be appropriate in the programming of certain subjects, such as neuroanatomy (Holland, 1960). Briefly, this consists of presenting a picture or diagram, the parts of which the student is required to identify. The first time this material is presented the various portions are identified. On subsequent presentations, more and more of the stimulus support identifying the various portions is removed. The first time a cross-section of the spinal cord is shown all the ascending and descending tracts are labeled. The second time the diagram is presented, the tracts are only indicated; arrows are drawn to them and perhaps the first letter of the name of each tract is indicated, followed by a number of blanks which may indicate the number of letters in the name of the tract. The third time the diagram is presented, the student would be given only the blanks to fill in. Fading can be done on an all-or-none basis—all the materials to be faded can be dropped out at the same time—or portions can be faded concurrently. Two or three of the tracts of the spinal cord might be named in the first presentation; in the second presentation, these might or might not be identified for the student, but other tracts would be faded. These diagrams might be alternated so that the number of responses required of the student on any particular frame might be kept quite small. He might not be required to identify more than four tracts per item. After he had identified a tract, on subsequent items that particular tract might be reidentified and others would be missing. In this way the material would constantly be kept before him, both in the stimulus form and in the form of repetitions of the required response.

A variant of the fading technique has been suggested for use with textual materials. This seems especially useful in converting the highly organized and information-packed data of tables into programmed form. As an example, in the construction of a program in neuroanatomy, we are confronted with tables listing twenty-two principal ascending and descending tracts in the spinal cord. Each of these tracts must be named; its origin must be stated; the student must know whether it consists of crossed or uncrossed fibers; its location within the cross-section of the spinal cord and in the longitudinal section of the spinal cord is important; its locus of termination and its function are important. One hundred and thirty-two pieces of information are called for that are relatively unrelated to one another and have a minimum of redundance. To introduce 132 separate items of information in conventional programming would itself result in a program of no mean proportions.

It may be that there is no recourse from such an almost prohibitive effort, but it is proposed that fading be tried with such material. We might program in the following way. The student would encounter the following statements: "The fasciculus gracilis has its origins in the dorsal root neurons below the midthoracic level. It consists of uncrossed fibers. It is located in the dorsal funiculus of the entire cord. It terminates in the nucleus gracilis of the medulla oblongata and its function is proprioception." The second time the student encountered this particular item in a frame the statements would consist of the following: "The fasciculus gracilis has its origins in the dorsal root neurons below the midthoracic level. It consists of uncrossed fibers and is located in the ______ ______ of the entire cord. It terminates

in the nucleus gracilis of the medulla oblongata and its function is proprioception." The third time the student encountered the statements he would read: "The fasciculus gracilis has its origins in the ——— ——— ——— below the midthoracic level. It consists of uncrossed fibers and is located in the dorsal funiculus of the entire cord. It terminates in the nucleus gracilis of the medulla oblongata. Its function is proprioception." The number of responses called for might be gradually increased until the final frame called for all five items of information concerning the fasciculus gracilis. The specific order in which information is dropped out is determined randomly. It is clearly undesirable to set up any patterning of dropping out. In a preliminary experiment it has been demonstrated that this procedure does succeed in teaching an appropriate verbal repertory in neuroanatomy. Whether this technique compares favorably with more conventional methods of programming is yet to be determined experimentally.

The fading technique lends itself admirably to any material, such as map reading, where visual discriminations are important. Map reading is one of the greatest problem areas in the training of soldiers for reconnaissance duties. Programs in map reading that use the fading technique might be an effective way of solving this problem for the military. In conventional map reading training, the only feedback that usually occurs is where the trainee either does or does not get lost. More economical ways of training must be possible.

Fading can be used to teach three-dimensional visualization of material. Students can be required to identify objects in their relationship to one another from many different vantage points, thus obtaining a concept

of the solid object which is otherwise difficult to teach. In map reading, the identification of standard map makers' symbols with aerial photographs offers the possibility of making the map more meaningful to the trainee in terms of objects with which he is familiar. The addition of color, of course, adds still another dimension that increases the immediacy and reality of visual experience.

The Uses of Rapid and Gradual Conditioning Techniques

It is suggested that the distribution of frames in subsections of programs be organized to take advantage of two processes that compete in their immediate effects, both of which are necessary for the establishment of an effective behavioral repertory. We know that conditioning is most rapid where continuous reinforcement is employed and where the material to be learned is grouped in a tightly organized mass. This technique of training or study rapidly increases response probability. We know it as "cramming." On the other hand, the maintenance of behavior over a long period of time requires a distribution of reinforcement, either by scheduling the reinforcement or by spreading training sessions over a longer period of time—perhaps even by conducting them in varying environments.

The mechanics of imposing a schedule of reinforcement upon a program are complex. It has been suggested that the program with a high error rate constitutes a program with an implicit schedule. It is said that the program with an error rate of 0 constitutes the 100 percent reinforcement condition of CRF (Carr, 1960). This is

correct. An *error* on the part of a rat undergoes nonreinforcement or is extinguished by CRF. This is directly analogous to the student's having made an error on a particular programmed item. The scheduling of reinforcement implies that the *correct* response has been made without reinforcement. The effectiveness of the schedule of reinforcement does not depend upon nonreinforcement of *inappropriate* behavior but upon the reinforcement of *appropriate* behavior. This is what increases the resistance of that behavior to extinction. Of course, one should never reinforce inappropriate behavior by CRF or anything else. Inappropriate behavior always goes unreinforced. The differential reinforcement of appropriate as against inappropriate behavior is what results in the differentiation of that behavior.

A combination of techniques might be employed to gain the benefit both of rapid conditioning and of the more effective maintenance of behavior through the spacing of conditioning procedures. Let us conceive of part of a program composed of four sections. These parts may be conceived of as individual items, although they might be larger units. Let us label these four sections A, B, C, and D. We shall employ two repetitions of each part of the program so that each subsection of the program is presented three times and two of its presentations will be consecutive. We might, for example, present part A followed by another presentation of part A. This would be followed by other sections of the program and, finally, by a third presentation of part A. Alternately, we might present part A, followed by other parts of the program, then a second presentation of part A, followed immediately by a third presentation of part A. The four subsections of the program first outlined might be pre-

sented as follows: A, A, B, C, D, A, D, B, etc. Such an ordering of presentations has several advantages where fading of textual materials is employed, as suggested for the neuroanatomy program. Bunching the material together might increase the likelihood that a student would be able to answer or complete the statement successfully, and spreading the repetitions of the items throughout the course of the program, interspersing them among one another, should increase retention of the material. This creates a somewhat different task for the student. Where repetitions are bunched together, the student has only to remember from one item to the next what needs to be filled in. The response that is reinforced may be under the control of very short-term contingencies. Stringing the items out through a series of frames requires a more demanding performance on the part of the student. He must respond to the material in such a way that the controlling variables are effective over a long period of time. Presumably the latter conditions would result in greater retention.

Repetition of particular items provides us with an opportunity to employ a technique we discussed in Chapter Five, namely, the scheduling of reinforcement. We cannot schedule reinforcement in the early stages of learning because the most effective acquisition of a response repertory during the differentiation phase is accomplished with continuous reinforcement. To impose a schedule of reinforcement upon the initial differentiation procedure might in the long run result in a more firmly established repertory, but it would decrease the speed with which the repertory was acquired. The spacing of reinforcement, although it contributes to resistance to extinction, also decreases the rate of initial con-

ditioning (Green, 1955). Therefore, if we were to make use of variable ratio reinforcement in a conventional program, it would almost have to be used only in review sections where material is presented for a second or third time.

Several researchers are exploring the possibilities of imposing variable ratio schedules of reinforcement upon programs. Skinner suggested this possibility (Edwards, 1956), and from the standpoint that it might be expected to increase retention, it is a logical development. Programmed instruction offers the possibility of making more extensive use of powerful laboratory techniques than is true of conventional instruction because the stimulus-response and response-reinforcement contingencies are explicit and are therefore amenable to manipulation. We could impose a variable ratio or VR schedule upon the acquisition of the verbal repertory almost from the beginning where the fading of textual materials is called for. This could be accomplished as follows. The first time that a term would be required as a response, the answer would be given by the program as usual. Perhaps the second time the response were required the answer would also be provided. But the third time the student encountered this particular response, the answer would be withheld. In other words, the student is presented the item, makes his response to it, and then finds that no answer is given. The next time, however, that this particular response is required, the answer will again be provided. At first, the proportion of frames wherein the answer is withheld is small. Gradually, as the concept or term is used repeatedly, the ratio of unreinforced responses might be increased. As this ratio is increased to 1.00 the program approximates an examination. At one extreme,

the presentation of a statement to the student is essentially the same as presenting material in a book—the student reads it. By altering the ratio of reinforced responses to the case where reinforcement is completely withheld, the program represents phases of instruction varying continuously from the form of a book to the form of an examination. Experimental analysis will establish whether this technique represents a useful adjunct to the techniques of programming.

Another use of vanishing or fading has been employed in the teaching of foreign language vocabulary. An interlinear translation of the text is provided. With the introduction of a new word, the English equivalent is written below it. The translated equivalent is provided for the first, say, five occurrences of the foreign word. After that, the foreign word appears without the English translation. This is a transition from the presence of the supporting term to its total absence. This transition can be made more gradual by dropping out letters in the English equivalent until by perhaps the fifth presentation of the foreign term, there is no longer an English stimulus accompanying it at all.

Inductive and Deductive Programming Techniques

Homme and Glaser have proposed a technique of analyzing programs in terms of items constituting rules or examples of the concept to be taught (Homme and Glaser, 1960). Is it more effective to proceed inductively

to rules through the accumulation of illustrative examples, or is it more efficient to state a rule and then prove it deductively by examples? This question is one that is argued extensively in the philosophy of science. How, in fact, do we learn about nature? Do we learn to generate general statements about our accumulated experience? Do we proceed inductively, or do we analyze the situation, form a hypothesis, deduce consequences of that hypothesis, and then test? We probably do both. We could not think rationally about a problem if we had not had the experience to recognize the existence of the problem. Conversely, if our efforts were devoted exclusively to the collection of facts, we could not deal as effectively with our environment as we might if we were to organize those facts into some kind of related structure.

For two reasons, a combination of inductive and deductive programs is probably most effective. First, a combination of the two enables us to proceed logically from one level of analysis to another. By combining techniques, we can more closely approximate the experience a student might have were he to encounter the problem in nature and solve it himself for the first time. Secondly, diversification of technique reduces somewhat the monotonous repetition of form that results from the monolithic adoption of one procedure. Gilbert makes the point again that the program one employs—the form of program, the use of prompts, panels, or other techniques—hinges upon what is most effective in the particular case. One adopts the procedure that works, recognizing the limitations of whatever technique it may be. If the required response is the end product the programmer de-

sires in his student's repertory, then that is the technique he should adopt. In reality, the programmer has no control over the technique he must adopt. If he is as sensitive to the behavior of the student as he should be, the student will determine the technique that the programmer must adopt. The programmer should have very little to say about the ideal form of his program before the student has taught him how to construct it.

The development of the simple linear program over the drop-out device was dictated by considerations of cost. The linear program is a more economical program with which to work. This is only one of the many important considerations that determine for the programmer the nature of his program.

Let us return for a moment to the question of optimal step size and its implication for the structure of a program. Skinner originally took the position that the smaller the step size the better. The error rate should be reduced to zero where possible. The optimal step size now is recognized to be a matter of experimental determination. It is a function of the ability level of the student population with which the program is to be employed.

Step size is an important property of a program. This property is the rate at which new concepts are introduced. A program with large step size introduces new material at a higher rate than a program with a small step size. It is more than reasonable to assume that step size is directly related to the error rate the student achieves with respect to the program. A technique has been devised to measure the step size of a program independently of the student's behavior.

Program Density

It has always been possible to measure the difficulty of a program through conventional item analysis. Difficulty level can be related to step size. Item analysis, however, is dependent upon the performance of the students. One cannot measure the difficulty level of an item without recourse to measurement of the behavior that the item calls forth. This behavior is subject to contamination by variables not under control of the programmer. Ideally, an independent measure should be used. We have used the type/token ratio as the measure of the density of a program (Green, 1961). The analysis is done as follows. A tally is made of the number of *different* responses required of the student in a section of a program. This number is then divided by the *total* number of responses required. A program would have a density of 1.00 if every response required by the program were different. The program would have minimal density if every response that is required of the student consisted of the same word. The tallies are of programmed answers, not *student* answers. There is an artifact in this ratio in that the value of the density ratio decreases as the size of the sample increases. Comparison, therefore, of two sections of a program must be made on samples of equal size. We have found that the density of a tape as we have defined it here, is significantly related to the error rate that students achieve with respect to that tape ($\rho = 0.47$; $\rho < .005$). The denser the tape, the higher the error rate.

We have calculated two types of density. The first we have called *independent* density. This is the density

of a single tape composing a part of a program. It is to independent density that error rate is significantly related. The second we have called *cumulative* density. Cumulative density takes into account the prior appearance of specific terms on preceding tapes. Plotting the cumulative density over tapes comprising subsections of a program gives a picture of the structure of the program that is useful for experimental purposes. As is seen in Figure 8, there is a sharp decrease in cumulative density within subsections of the program in medical parasitology. These decreases represent the exhaustion of the special technical vocabularies relative to each subcategory of parasite. Superimposed upon these curves is a gradual decrease in the initial cumulative density of each subsection of the program. This more gradual decrease describes the rate at which the general terminology is exhausted by the program. With such an independent specification of the density of the program, one can arbitrarily manipulate the program to control the rate at which concepts are introduced. Experiments have been performed upon the effect of changing step size in programs (Coulson and Silberman, 1960). Independent density and cumulative density provide ways to specify precisely the composition of a program in this regard. One can now equate programs constructed by different techniques in terms of the rate at which they introduce new material.

The density function is an indirect measure of the rate at which material is introduced. Material occurs in the body of frames that is not directly represented by the density measure. However, examination of programs to which this measure has been applied reveals that a greater density in the answers parallels greater repeti-

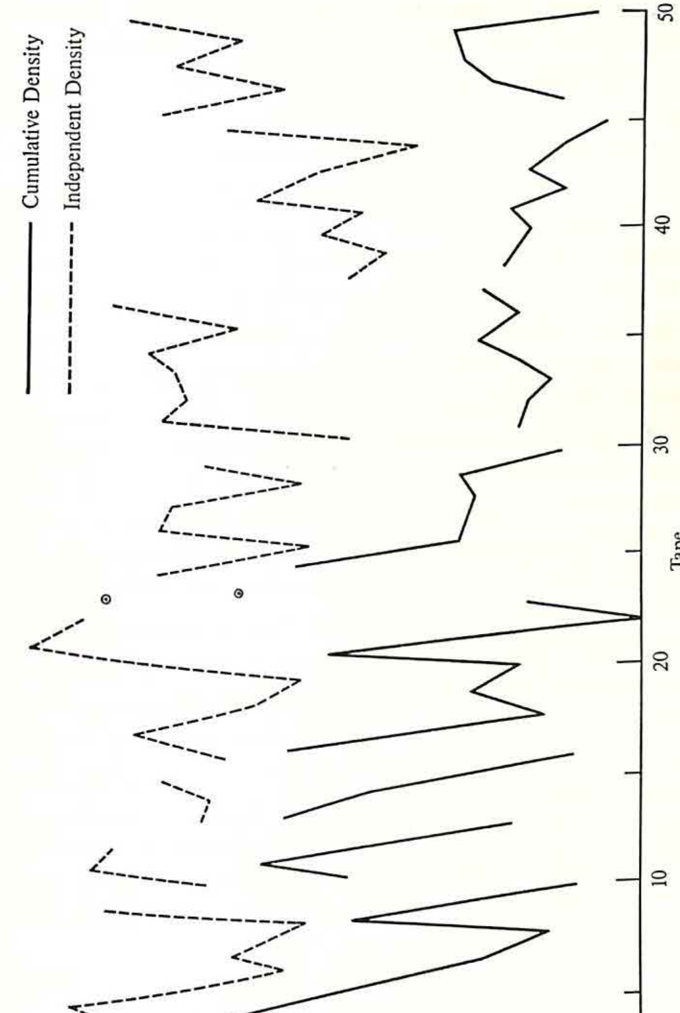

FIGURE 8. Densities by parts of program in medical parasitology.

tion of that material within the frames themselves. It does provide an indication of the degree to which material is weighted or repeated, even though the indication is indirect.

A suggested use to which the density function might be put in research is the assessment of the tolerance of a given population of students to the density of programmed instruction. A particular group will have a higher tolerance, and programs can be designed in advance for these populations. Parallel programs can be constructed covering the same material with lesser density. We do not propose that the specification of program density is a substitute for extensive evaluation, but time might be saved in the arduous task of program writing if the programmer had some notion that his program was approximately within a range that was appropriate to his students.

An additional advantage of independent density is the insight it gives into the behavior of the programmer. As mentioned before, one of the most difficult tendencies against which the neophyte programmer must guard is the tendency to write examination items rather than programs. Inspection of independent density of program samples produced by various programmers shows this tendency in that the programmer's initial work shows a higher independent density than does the later work of the same programmer. As programming skills are acquired, independent densities decrease. A decrease in cumulative density reflects the exhaustion of the specialized vocabulary of the subject matter; a decrease in independent density reflects the acquisition of programming skill by the programmer.

Difficulty Level and Performance Variability

As we have seen, the density of a program is significantly related to the error rate achieved by students working on that program. What can we predict concerning other behavioral effects of programs of varying difficulty levels? Let us suppose that the probability of a correct response is inversely related to the difficulty of a program. The program can produce probabilities of correct responses varying between two extremes: where the probability of a correct response equals 1.00, and where the probability of a correct response equals 0.00. Suppose also that the discriminative stimulus is composed of elements and where the response probability is equal to 1.00, the response is conditioned to all the elements. Where the probability of response is equal to 0.00, the response is not conditioned to any elements. Let us suppose that the organism samples only a proportion of all the elements at a given time. Where response probability is not equal either to 1.00 or to 0.00, the discriminative stimulus is composed of some elements that are conditioned and some that are not. On a given trial, the student responds to a subset of all the possible elements composing the totality of the discriminative stimulus. Some of these elements will be conditioned, some of them will not. Let us assume that on the average, the same proportion of elements will be conditioned within the subset as within the total field of elements. On a given trial, the subject may fortuitously only attend to those elements which are conditioned. He may respond on other occasions to elements none of which are conditioned. Response proba-

bility, then, is unstable from trial to trial, even though response probability may theoretically be defined as the proportion of conditioned elements to the total set. It will vary between the values of 1.00 and 0.00 from trial to trial. It can be shown that as probability of response deviates from the extreme values, either 1.00 or 0.00, the student's behavior becomes more variable. This can be shown to follow mathematically from the theoretical variance of probability of sampling or discriminating subportions of the stimulus field. Experimental evidence has been gathered in the context of simple visual discrimination that supports this prediction.

If the error rate of a program increases in proportion to the unconditioned stimuli in the situation, then the variance in the error rate should be maximal, somewhere between an error rate of 0.00 and an error rate of 1.00. If the mean error rate of a program is 0.00 or 1.00, there can be no variance at all. It is almost trivial to say that error rate will be maximal between these extremes. Table 1 shows the variances in error rate obtained from the seven most difficult tapes and the variances in error rate taken from the seven tapes with the lowest mean error rate. The F ratio of the two variances is highly significant. As the error rate increases, variance in performance increases.

TABLE 1. Means and Variances in Error Rates for the Seven Most and Seven Least Difficult Tapes in a Program in Medical Parasitology

Difficulty	N	Mean Error	Variance	F	p
Greatest	7	.106	.0374	5.06	$<.001$
Least	7	.024	.0074		

A second predication concerning the variability of behavior has to do with the abilities of the student population. In another discrimination experiment it was found that slower subjects who learned at a lower rate also gave a more variable performance. Analysis of data drawn from our programs shows the same relationship in the behavior of medical students responding to a teaching machine program. This is but one more piece of evidence in support of the argument that the processes of the learning laboratory are the processes of the classroom.

Program Evaluation

It has been repeatedly said that programmed instruction offers greater possibility of the close monitoring of students' behavior than any other teaching technique. Because of this, it is a more flexible instrument. We have said that the training of the programmer depends upon the reaction of the student to the program. In concrete terms, just how is all of this brought about? In terms of the evaluation. There are two ways by which this evaluation is accomplished. In the first, the programmer is face to face with a single student. The student works through the program and the program is adjusted on the spot in response to his errors. New frames are inserted or ambiguities are cleared up. The second is the classroom test of the program. Here the answers are recovered by the programmer and subjected to analysis. This analysis is not simply an item count but involves the precise incorrect response. Not only is there feedback to the programmer telling him that students miss particular items, but there is feedback telling him what errors the students

made. The programmer has the ultimate critic's evaluation before him before the program ever goes to print. For this reason the program is more flexible than the textbook. Not only does the programmer know that a particular item is bad for one reason or another, but he knows why it is bad.

Often a student responds successfully to a series of frames, but at some frame his error rate increases sharply. The failure of the program may involve one of two things. The programmer may have assumed too much on the part of the student, and at that point the continuity of the program is broken. It is broken because too large a step has been taken. Error rate may thereby be increased. Or the particular item may have been worded ambiguously. Once the ambiguity of a particular frame is cleared up, the rest of the program may operate smoothly. The onset of a higher error rate does not necessarily imply a gap in programming. Often an ambiguity within a single frame will have a cumulative effect of increasing the error rate. This is particularly true where that particular frame introduces a new concept, one that is essential to a whole series of succeeding items. Often the single item that is giving difficulty requires a change in but a single word or a change in the response required by the item. Occasionally the item may need to be split into two. Too much information may have been given in an item for the student to digest and use. The frame itself may simply contain excess information. One school of thought dictates that a frame should contain no excessive information whatsoever. The only information that should be in the frame is that required to answer the particular question or to complete the particular statement. Another factor that may cause high

error rate is demanding too much in response. The programmer may have required too many answers to a particular frame. When an excessive number of answers are required the linguistic structure of the frame is demolished. A student can do very little with a statement that is composed primarily of blanks. There are simply too many degrees of freedom.

The technique of programming is developing rapidly. New approaches are continually being devised. We have just begun to explore the possibilities of these techniques. It is folly to assume that what appears to be the most effective technique today will remain so for long. Excellent manuals are continually being developed to bring new procedures to the programmer. All that can be said in summary is that the best techniques must develop from the application of our knowledge of the processes as they have been explained here. If the programmer understands the nature of his experimental organism and remains sensitive to his needs, listens to his advice, he cannot help but construct effective programs.

CHAPTER NINE

Evaluation

The evaluation of a teaching machine program begins with its construction. Persons experienced in programming agree upon designing the program around the behavior of the student as the program is written. This is using the subject as his own control. The programmer builds his program on the developing repertory of the student. This technique has been employed extensively in the extensive studies of the behavior of infrahuman organisms made by Skinner and others. Skinner has said of this procedure (Skinner, 1958):

> Research of this type is usually single-organism research. Any other experimental method is often impossible. When an experiment on one pigeon runs to thousands of hours it cannot be repeated on even a modest group of, say, ten subjects, at least if one wants to get on with other matters. Fortunately a statistical program is *unnecessary*. Most of what we know about the effects of complex schedules of reinforcement has been learned in a series of discoveries, no one of which could have been proved to the satisfaction of a student in Statistics A. Moreover a statistical approach is just *wrong*. The curves we get

cannot be averaged or otherwise smoothed without destroying properties which we know to be of first importance.

Gilbert has also stated the position in the present context (Gilbert, 1960):

> The first characteristic of the human operant conditioning laboratory is the provision for intensive, long-term study of a single human in an artificially and fully controlled environment. This logic is based in the universal agreement that the human comes to the laboratory with a complex conditioning history. The traditional approach has attempted to get behind idiosyncracy by statistical averages of short behavior samples. What this gives us instead of generalizable and timeless principles is an average of cultural effects. The traditional laboratory is too artificial for an interesting sociology, too sociologically sensitive for a culture-free "diachronic" behaviorism. We make each subject his own control and we can do this by virtue of the long-term investigations. We collect months of base-line data while standardizing and making unique the experimental space. A single subject may work in this laboratory for five or more years.

The Duration of the Conditioning Procedure

When we undertake to make a significant change in the behavior of a human being we cannot base our principles upon studies which are trivial events in the lives of the experimental organisms. Experience in the animal

laboratory shows, as no exposition in experimental design or laboratory method could possibly show, the importance of pursuing the study of a single organism for a long period of time. In one study a mixed schedule of reinforcement was being used. For 6 weeks two birds were subjected to the particular reinforcing contingencies of this study for 8 hours a day. Within 24 hours of one another, both birds began to give very peculiar performances. They would respond at an extremely high rate that was steady and unvarying, or this behavior would break and they would respond at a rate about one-fourth of the higher rate. The slower rate was also steady. This pattern of behavior continued for 2 or 3 days. Observing these animals in the experimental chamber revealed that the behavior that produced these records consisted either of rapidly pecking the key, or of pecking the key once, making swiping motions off to one side of the box and then returning to peck the key again, making more swiping motions to the side of the box and returning to peck the key, and so on. This, however, was not the final state of behavior. The organisms ultimately settled down to behavior wherein the sharp divergences in recorded rate of response did not exist, and where transition from one phase of behavior to another was more gradual.

Had this study not been pursued as it was we would never have observed the rather peculiar phenomenon that emerged after 6 weeks. Furthermore, had we stopped there we would not have known that this performance was not the steady state of behavior under those conditions. The temptation was great, it must be confessed, to terminate the experiment when the rather

startling divergences in rate were observed and to write a report on this remarkable effect of the mixed schedule. Fortunately the temptation was overcome, the experiment was continued, and more was learned about the behavior of these organisms.

A time sample of 1 hour or 8 hours or 2 days is not a very large period in the life of a human organism upon which to base broad general principles of behavior. A short learning task, whether in the context of the psychological laboratory or in that of the classroom, is less likely to reflect basic processes than to reflect momentary dispositions or indispositions of the subject organism. The specific changes in the behavior of individuals can be observed only as they occur. They are best observed when the learning is new, when the differentiation of that behavior is on-going.

In addition to the individual student evaluation of the program during its construction, the field test of a program is also an important and necessary mode of testing. We considered the individual evaluation process in the last chapter. The field test assumes even greater importance when it becomes impossible to pursue the individual evaluation of a program as one would wish. It is to the particular problems of such a test that we shall address ourselves in this chapter. The modification of a program to accommodate the behavior of the single organism is simple and direct. It is dictated by the conditions of the moment-to-moment interactions of programmer and student. Because the classroom evaluation is less adequate, many precautionary measures must be undertaken.

Learning consumes time. The experimental evalua-

tion of programmed instruction plunges us into all the problems of experimental design related to time-ordered phenomena. Unfortunately for statisticians, the components of the behavior of organisms are not independently organized the way onions and asparagus are bedded in different parts of a garden. What goes before helps to determine what comes after. One can make use of the prior experience of an organism and build upon the learning that has taken place.

In using the organism as its own control, one takes advantage of the time-ordered property of behavior. One observes behavior for a long enough time to determine some sort of base-line performance. Then the experimental, or, in this case, training procedure is introduced. Behavior following this is compared with base-line behavior to evaluate the effects of the procedure. The procedure may then be terminated and the behavior subsequent to its termination compared with behavior under both the prior conditions. The nonindependence of the organism's behavior over time is an asset. On the other hand, this nonindependence constitutes a serious methodological obstacle from the point of view of statistical control. Several good textbooks deal with techniques of experimental design, problems of counterbalancing, and so on. These go into much more detail with respect to the peculiar time-order problems than we can do or wish to do in this treatment. The reader is referred to these studies in methodology. Let us in the meantime examine some of the specific problems with which the programmer will become involved in the course of his evaluation, irrespective of which experimental design he may or may not employ.

The Test Population

The programmer's first concern is with his test population. The teacher who seeks to apply these techniques in his classroom should have no such problem. He should know his students and their state of knowledge. The programmer who works at some distance from the target population needs to inform himself carefully of the composition of the population. A knowledge of sampling theory is needed in testing or evaluating a program to insure that the programmer is dealing with a sample that is representative of the population with which the program is ultimately to be employed. The background or preparation of the target population to work with a program is very important. A programmer obviously should not develop materials requiring advanced reading ability for a group of kindergarten children. Carr has listed as one of the major sources of error in programmed instruction the erroneous assumption that students are prepared to deal with the initial steps of the program (Carr, 1961). Most programs developed for primary and secondary school and college purposes are based on the assumption that the student possesses sufficient reading skills to enable him to respond intelligently to the program as a program. Where this assumption is unwarranted, the program will not accomplish anything. The same problem exists where the program begins at a more sophisticated level. Programs proceed from the known to the unknown. It cannot be emphasized too strongly that the initial steps of a program do indeed have to do with the known rather than the unknown.

The ultimate criterion of evaluation must be the

performance of students on some criterion measure or instrument. The programmer should be familiar with conventional procedures of test construction and validation. Where the evaluation of an educational technique is at stake, these issues assume greater importance than they do in the conventional teaching context. In an ordinary examination the educational system is not being challenged. The evaluation of a program places the burden upon the program rather than upon the student. This change of emphasis makes aspects of testing that are not always evident pertinent in relation to the teaching process.

Perhaps it would be informative to illustrate some of the problems of evaluation in terms of an experience with a program in the medical sciences with which the writer was associated. The evaluation undertaken was not intended as a serious research effort but arose fortuitously as a consequence of the unexpected availability of test data from previous years. Our early experience in programming, like the experience of so many others, is instructive primarily because of the many things that it has taught us not to do.

The first problem we faced was the lack of sophistication in the programmers. None of the programmers engaged in this project were familiar with the technical aspects of the subject matter to be programmed. A set of textbooks and reference material were available. We were provided with lecture notes that described in broad outline the subject matter to be covered. Available to us also was the assistance of the expert teaching the particular course. His time, however, was limited by demands from other sources. His relationship to the program finally became that of review editor, as he would

examine the program as it was developed for technical accuracies. A program of over 2,000 frames eventually was written from textual material available to us. The obvious limitations of the programmers in their sophistication with respect to this particular subject matter made inevitable a divergence in the emphases given various portions of the program compared with the emphases which the students desired, as given by the lecturer. Despite these not inconsiderable problems, an initial program was constructed and used experimentally in 1961.

One objective of programmed instruction in the medical sciences is to release students from an already overburdened schedule so that they may devote more time to research activities. Because we were working in this context, it was impossible for us to use students in the initial composition of the program. The first test of the program had to be a field test. Students were requested to make use of the program as a form of supplementary instruction but they were not required to do so. Consequently, only twelve out of a class of twenty-three students actually completed the program. A questionnaire given at the end of the course revealed motivational problems, morale problems over and above the natural consequences of testing an untried program. Students were heavily pushed for time. The students were not given free time to work on the program but had to do it on their own study time. Some of the students, particularly those who lived farther from the campus, found it particularly difficult to work on the machines. Some of this negative reaction might have been avoided had the students been able to work in their dormitory rooms either on machines or on programmed textbooks.

We had intended to administer the first draft of this program simply to gather information concerning error rate for a redraft. However, the professor in charge of the course gave the same final examination as had been given the preceding year. This enabled us to gather more information than we had originally intended, but it was information of such a tenuous nature that it illustrated, among other things, most of the things that can go wrong with experimental evaluation of this type. The most encouraging result was that the mean examination performance of the class using the program was over 10 percent higher than the mean performance of the preceding class, which had not used the program. What does such a result imply? Naive enthusiasm might lead one to conclude that the program produced this difference. However, the samples involved in both classes were small, not over twenty-four students in each. Further, selection procedures and the general level of students applying to a particular medical school can vary from year to year. With the increase in number of applicants, an effort is made to select better students every year. The difference in performance in favor of this year's group over last year's might be directly attributable to more effective admission policies. In view of this, there is little unqualified enthusiasm for attributing the performance difference to the program.

Another fact is that the twelve students who completed the program came from the lower half of the class in terms of their performance in other academic subjects. The students who did not complete the program were generally the better students. In other words, those who completed the program needed to complete the program. A selective factor, therefore, biased the per-

formance of the part of the class that completed the program against greater success on the final examination. As it turned out, those who did complete the program were exactly as successful on the examination as those who did not complete the program. Proving the null hypothesis is impossible, but the pitfalls of arguing that the program helped the poorer students because there was no difference between the groups should be obvious. The same effect could be produced by an examination that simply did not discriminate between good and poor students as defined by their standing in other courses.

Correlation between the number of tapes completed and performance on the final examination yielded a slightly, though not significantly, negative relationship. This negative correlation in favor of those who completed fewer programs is attributable, at least in part, to the fact that the better students by and large did not complete the program. No programmer wishes to consider the equally reasonable hypothesis that the program is injurious to performance. Adding to the despair of programmers was the correlation between number of tapes completed and performance on the final examination for those students who did not complete all programs. This correlation was also negative. However, the class standing of this latter group showed a high positive correlation with their performance on the final examination. Again, the program was working against the selective bias that operated over-all in the class; namely, that the less able student completed more tapes.

Relevant to our concerns with motivation was the resentment engendered in some of the more able students by the concept of programmed instruction of medical subjects. At least one of these students avowedly set

himself the task of demonstrating that he could outperform anyone who had completed all the tapes. Of course he succeeded. His success, however, must be laid to the fact that he was competing with the program. To the extent that his performance helped to reverse the anticipated results, perhaps his results should be claimed as one of the effects of programmed instruction. The attitude questionnaires that were given at the end of the trial indicate that several students reacted in this way to some extent.

Comparing Programmed Instruction and Classroom Techniques

To what extent may programmed instruction be compared with existing techniques? In the trial that we concluded we did not compare programmed instruction with conventional techniques of instruction. Rather, we compared the performance of students who used the program in addition to conventional instruction with the performance of students who had only conventional instruction in a preceding year. We also could compare the behavior of students who had more programmed experience superimposed upon their conventional instruction with those who had less. We would hazard the opinion that in most practical situations it will seldom be possible for a pure programmed treatment to be compared with a pure nonprogrammed treatment. It is not likely that one can test technique A against technique C. It is more likely that one ends by testing technique A + B against B + C. Even if it is practical to compare A with C, good experimental design demands that the amount of time

spent, the difficulty of the material, or some independent property of the instructional process be equated between the two techniques. Where one technique consists of conventional classroom teaching and reading assignments, this is extremely difficult to achieve. One might argue that it is logically impossible to equate techniques in such a meaningful way. Barring the development of new techniques of evaluation, we must be content for the present with compromises in experimental tests of the classroom effectiveness of programmed instruction.

One of the most elegantly simple approaches to this problem is that used by Holland and Skinner in their studies of the program in the *Analysis of Behavior* at Harvard. The technique of evaluation involved categorizing items of the examination in terms of their relevance to the program. There were items that were directly relevant to the program; that is, certain parts of the examination covered material explicitly dealt with in the program. Another set of items were indirectly relevant in that these parts of the examination were generally and broadly dealt with in the program. Finally, there were those items which were irrelevant to the program. The program did not deal with material covered by those parts of the examination. The performance of students on those items which were directly relevant to the program decreased almost linearly as an inverse function of the number of tapes completed by the students. The more tapes completed, the higher the performance on these particular items. The same was true to a lesser extent of items that were indirectly relevant to the program. If only these two items of information were available the results would be confounded with motivational variables. The student who completed more items might

be more highly motivated and therefore more likely to perform well on an examination anyway. Although this is in contradiction to our findings in the medical school setting, it remains a reasonable hypothesis.

The design of the evaluation procedure provides an answer to this objection. The performance of these students on items that were *irrelevant* did not change as a function of the number of tapes completed. This design, using the students as their own control, yielded information that would have been extremely difficult to get and whose interpretation would have been extremely tenuous to make by traditional statistical designs involving classical control and experimental groups. The fact that the responses to the three categories of items in the examination were made by the same individuals strengthened the findings; had factors other than the program been responsible for the results, performances in all three categories would have been highly correlated. That they were not strengthens the validity of the findings because the results were contrary to alternative hypotheses. This is an example where the nonindependence of behavior within the same subjects, using the subjects as their own control, is a more effective technique than more conventional experimental designs.

Determinants of the Error Rate

To return to our experience in evaluation, what about the criterion? What does performance on a particular test mean? Our analysis illustrates additional problems of interpretation that may be encountered. We analyzed sections of the examination dealing with specific parts

of the program. Counts were made of the proportion of errors on the examination as they related to the error rate of corresponding parts of the program. We found no consistent relationship at all between these. Such a relationship, had it been found, might have indicated that the more difficult the program, the more—or the less—did it contribute to criterion performance by conditioning appropriate behavior. Fortunately no such relationship emerged. It was fortunate because in the absence of such a relationship, more penetrating analysis was required than would have been necessary had it been obtained. It is not unimportant to account for failure to establish results. It is important to know that results can be produced in terms of alternative variables—especially when some explanations are less obvious than others.

What contributes to error on an examination? This is the basic question that defines the role of examination in relation to programs and their evaluation. One thing that should determine error on an examination is the effectiveness of teaching. The better program should presumably decrease the likelihood of error. The term "better," however, is relative, for some materials are intrinsically more difficult to program than others. They are intrinsically more difficult for a number of reasons. One reason might be the necessity for supplementary devices of presentation, recourse to laboratory demonstration and the like, which might not be true of another course. Moreover, one subject matter may be more difficult, not only to program, but simply more difficult. Concepts may be more abstract; there may be a larger number of definitions; there may be a larger body of knowledge to be communicated to the student than in another subject. Differences in the ease of programming, differ-

ences in the basic underlying difficulty of the material, contribute to the rate of error on the examination. The ability of the student, independent of his training, will contribute to the error rate. His ability is, in turn, dependent on his motivation, his innate limitations, the background of preparation he brings to the course in the first place. All of these factors contribute to the error rate. Last, though certainly not least, is the examination itself. Is the item that occasions the error an item that can reasonably be answered by someone with a basic grasp of the material which it seeks to examine? Is the item poorly written? Is it ambiguous? Does the item bear any relationship to competence in the subject matter? Or is it, in terms of item analysis, a poor discriminator? A poor discriminator is an item that introduces a wider range of possible variables controlling behavior than are required for the correct response. Item analysis cannot give us a direct indication of either the difficulty level or the discriminability of an item that is uncontaminated by other variables. Conventional item analysis is based upon the response of students to that item and, as such, necessarily incorporates all the other controlling variables that contribute to error on a response. Item analysis, then, does not escape the problem but in some circumstances even compounds it. For this reason, independent description of a program in terms of its density has a considerable advantage over the simple count of the errors that students make to items of that program. Density is independent of the student's behavior. Fortunately, it is also known to be related in a meaningful way to that behavior. It is related to, *but not defined by,* that behavior.

A significant negative correlation was obtained be-

tween error rate and performance on the final examination. We shall not take this as evidence that a high error rate produces poorer learning, although our particular systematic orientation might tempt us to do so. One must approach with caution the results of correlational analyses. What does the correlation coefficient imply in this case? There is only one thing of which we may relieve ourselves with respect to correlation in this instance: the interpretation that performance on the examination determined the error rate on the tapes. This interpretation of correlation is the only one that may be disregarded. Another interpretation is that the less able student performs less adequately on any task he is given. Yet another interpretation is that nothing was learned whatsoever in this particular experiment, and that the questions which were asked in the examination and the material covered in the program dealt with essentially the same material. People may have been uniformly ignorant and others uniformly well informed by virtue of prior learning experiences. This is a disheartening but possible interpretation. Correlation has a place in evaluating teaching machine programs, but it must remain a minor place. The ultimate analysis of any technique of behavioral control must rest upon the establishment of functional relationships rather than upon the computation of correlation coefficients.

Although these problems were exaggerated in the uncontrolled context in which we encountered them, they exist in every serious evaluation. They are the problems of extraneous control and imprecise definition that are implicit in any testing. Parallel problems are found in the animal laboratory. That human learning and its analysis is accomplished under conditions of less specificity compounds the normal problems of laboratory ex-

perimentation. Definition of the test in terms of its relevant properties is the same order of task as the precise physical description of a lever in the rat box. The test is distinguished from the lever only in number of relevant properties and in the degree to which the experimenter can control or specify them.

The Method of Evaluation

There are two problems regarding criteria that we encounter in evaluation. Is the criterion measure reliable? Is it valid? Is the test related in any meaningful way to the ultimate task for which the student is trained? It is not inconceivable that one may occasionally have more confidence in the programmed instrument than in the instrument of evaluation. The careful validation and standardization of a test requires considerable expenditure of time and effort. It requires the cooperation of potential users from a large and representative sample of the population. Such a program of test validation is unrealistic for the average teacher undertaking to program some of his material. What should such a person settle for in a test? The practical answer is to use the same kind of test he has been using for other methods of teaching. If his program is developed for commercial distribution, not only large-scale test validation but also a large-scale test of the program become necessary.

The intrinsic difficulty of programming some materials contributes to error on examination items. This point is important enough to justify some elaboration. A functional analysis of behavior allows one to study private experience. Private events, although directly acces-

sible only to the individual, differ in no essential way from more public events. This particular problem arises again because of this question of accessibility. It is theoretically possible to develop a program to teach discriminations of taste and to make discriminations of events internal to the organism. The monitoring of these events for appropriate reinforcement is very difficult. It is for this reason that discriminations of visual, auditory, and tactile sensation are sharper, more accurate than the discrimination of the location of a pain in the intestines. The stimulus material is available to the learning organism for accurate discrimination of internal events, but the verbal community responsible for the differential reinforcement of the correct responses with respect to those stimuli does not have access to information of sufficient precision to enable it to sharpen these discriminations. The inaccessibility of private material sets a limit on programming, as upon any form of training procedure where such material is involved.

In summary, program evaluation has two objectives. One is to determine whether the program teaches. The other is to determine how well it does the job. The evaluation of the program with individual students during its construction is the method of choice for attaining the first objective. The field test may also attain this first objective and may be useful in casting the program in its final form. The field test is inevitably involved in attaining the second objective. A field test really amounts to a full-scale experiment. It is a job of such complexity that its proper management should only rest upon someone professionally equipped. Not only are questions of statistical design involved, but a familiarity with the techniques and methodologies of conditioning is demanded. When a

program reaches the prepublication stage, it demands more than amateur competence. This is not to say that experimental psychologists who have conditioned rats are the only persons competent to evaluate programs, but a professional competence based upon recognition of the complexities of the task and a familiarity with appropriate analytic techniques is required. It must be emphasized that a knowledge of statistical methodology does not suffice. Statistics do not tell a programmer what to do next.

CHAPTER TEN

Problems

Formal education has only begun to feel the impact of the technological revolution set in motion by techniques of programmed instruction. The teaching machine and the program for the teaching machine are only opening guns in what promises to be a massive assult upon traditional concepts and methods of instruction. As an academician, I have long been fascinated by the ease with which administrators in the field of education make decisions and pronouncements concerning ideal ways of achieving educational goals without the least fundamental acquaintance with the facts of the animal behavior laboratory. It is only within the past decade that behavioral control techniques have been applied to the practical problem of education. The advent of the teaching machine itself is an event of great import. Of greater significance is the implication that it has upon the entire theory of education. The most significant and difficult problems—problems we cannot yet anticipate—will stem from the broader applications of behavioral control techniques. Certain problems, however, can be anticipated. These problems may be categorized as problems of pro-

gramming, problems of production and the economics associated therewith, problems of school administration, problems of effectiveness, and problems created for the student.

Production of Machines

The production of devices with which programs may be used far outstrips the production of actual programs for such devices. The demand of students and schools for programs they can use is large and growing. The generation of good programs is a time-consuming and expensive process. The temptation exists for the entrepreneur to throw hastily together programs which will sell but which will not accomplish the objectives they purport to accomplish. That this type of exploitation is inevitable is unfortunate. As an answer to this problem, the general public must be educated to the objectives and techniques of programmed instruction. Particularly, they should be familiar with the statement regarding criteria for programmed materials that was released by a Joint Committee of the American Educational Research Association, the American Psychological Association, and the Department of Audio-visual Instruction of the National Education Association. This statement is included as an appendix to the present volume. A published program should indicate the extent to which revision has been based on student response and preliminary tryouts and should indicate the nature and extent of classroom tests of each program. This would be some safeguard against the type of exploitation we have so pessimistically envisioned.

The Form of the Material

In what form should programmed material be published? At present, certain factions argue in favor of teaching machine programs. There are definite advantages to the controlled sequential presentation via the machine. Comparing the effects of the teaching machine and the programmed textbook, some studies have revealed that there is no difference in the performance achieved. For certain reasons these results are less than totally convincing. First of all, because the control of student response in the programmed textbook is less than it might ideally be.

One problem in the use of programmed textbooks is that the student is able to get cues for answering a given frame from his own answers if they are continuously before him. The student working through a program on a machine does not have this additional support. The student working through the programmed textbook presumably does have this support available if he writes his answers on a sheet of scratch paper. Such supplemental information increases the probability of correct response and possibly decreases the learning that can be acquired by working through the material. This incidental prompting tends in the same direction as excessive support given by unnecessary prompts. It is too bad that a program should be carefully constructed with all unnecessary prompts removed and then be cast into a format that permits the adding of such prompts to the situation by the student's own behavior. This is not a fatal flaw of the programmed textbook, but certainly diminishes its value in contrast to the machine. It may be more

than balanced by the convenience to the student of having his material continuously available to him. The particular machines that have been used in experiments comparing machines with programmed textbooks have suffered abysmally from mechanical breakdowns. In a real sense, these tests have not been fair tests of machines. Machines will not always be beset by such severe mechanical difficulties. Judgment had best be withheld concerning the desirability or effectiveness of the two techniques until machinery is available to make a legitimate test.

In support of the programmed textbook method of publication are two arguments. One is economic. Books cost less than most machines. In the absence of an adequate budget, a school system is unlikely to be able to purchase machines of any degree of complexity for use with programs. It may be that the most efficient method from the standpoint of such school systems would be the programmed textbook. The machine places the financial burden upon the school. The programmed textbook places the financial burden upon the student. Great advances have been made in the development of production techniques, which have reduced the cost of programmed textbooks from as high as $15 a volume to a fifth of that value. The development of plastic and wood machines has reduced what is essentially the same device from a cost of $80 per machine to a cost of $5 per machine. Machines of the latter category may solve the problem of the school system in purchasing machines. If machines of this type are found to be adequate, the burden for their purchase may be placed on the student. A program for a machine can be used over and over again by several students. If 10 students use such a program

and if those same 10 students would otherwise have purchased 10 programmed textbooks, at, say, $5 apiece, what must the program cost?

School Administration

Insofar as problems of school administration are concerned, the most immediate effect of widespread adoption of these techniques will be in the area of curriculum planning. The entire American school system is predicated upon a system of organization of material in curricular sequence to correspond with what are presumed to be the abilities represented by the age and grade levels. A pair of experiments give us a faint glimmer of what havoc must eventually be wreaked upon this entire superstructure of administration by the application of these techniques of behavioral control.

Moore has demonstrated that he can teach three-year-old children to read and write using specially constructed typewriters (Moore and Andrews, 1959). Komoski and Eigen have demonstrated that children in the primary grades can successfully be taught logic (Eigen and Komoski, 1960). What happens to concepts of reading readiness and maturity as they have so elaborately been developed in the face of such evidence? With the development of these techniques and others which are sure to follow, we must admit we know nothing concerning the ideal sequence of programming of curriculum that should be used. The greatest problem is not the challenge but the tremendous amount of work that will necessarily be involved in restructuring our educational system. Not only the practical structure but also the

basic concepts of our educational system will be revised by these events.

What happens when we learn, for example, that students may be capable of mastering ten years of formal education in half the time? The self-pacing feature of the techniques that have been developed will serve to widen the gap between the more gifted student and the less gifted. The problems of ability grouping will become ever more acute. These problems must never be allowed to dictate a rejection of these new techniques. Such a rejection can be predicated only upon a vested interest in the status quo. The exploitation of techniques of programmed instruction does not rest upon the isolated laboratory psychologist or the research group in a large school system; it rests upon every teacher and school administrator in the country. These techniques have existed for thirty years in the laboratory. In the laboratory they have contributed nothing to the advancement of education. Their use now depends upon their effective employment in the schools.

The Question of Motivation

What is it that contributes to the present effectiveness of these techniques? Is it, as we have argued, that they represent well-known powerful techniques that have long been used and studied in the laboratory, or is it that these techniques are just another "gimmick," another gadget with transient interest and popularity whose effectiveness will decrease with familiarity? Many persons have argued that the teaching machine depends for its effectiveness as a motivating device upon its resem-

blance to gadgets and toys. It has much of the charm of the pinball machine. Opponents argue that once students have become familiar with the device and have used it in a number of courses, it will no longer be as effective as it now appears to be. This is a reasonable and possibly correct objection. The only answer to it at this time is that although boredom does seem to increase as the student works through a program, there is no evidence that this loss of fascination does away with the effectiveness of the teaching device.

Effectiveness must be considered in a broader context in relation to curriculum design. At what level may material most effectively be introduced to the student? We no longer need concern ourselves with the question of the level at which it is possible to introduce new materials to the student; it is possible to introduce any material at any age to a student. It is likely, however, that there is an optimal sequence that will yield the greatest effectiveness. We shall, within the next few years, begin to see the results of studies directed toward answering this question.

What is the most effective role of the teaching machine? Writers, including the present one, are prone to make glib statements about the challenge presented to the teacher by the new technique. We say that programmed instruction can free the teacher to pursue more challenging and stimulating aspects of teaching because he no longer needs to devote his efforts to getting across basic facts. When we speak of a programmed presentation of the basic substantive material of a course, what do we mean? What facts constitute basic facts? How large should a program be? What percentage of a student's time should be devoted to work on the program

in order to achieve the most effective integration of programmed instruction with lecture and demonstration? At present, the answers to these questions remain experimental.

Problems for the Student

What problems are created for the student by programmed instruction? One problem will be an exaggeration of the problem of age grouping versus ability grouping. The bright student is now faced with the prospects of staying with a group that is chronologically of his generation or of moving into an older age group whose abilities, knowledge, and sophistication are commensurate with his own. This situation will be aggravated. Perhaps too much emphasis has been given the emotional disturbances or potential emotional disturbances resulting from moving a gifted student ahead. The problem will certainly become more acute, and it will be a problem for the gifted student to solve. On the other hand, the problems of slower students' falling behind may also be aggravated. On the positive side, the self-pacing feature of the technique makes it possible for a student to make up missed materials efficiently and with little expenditure of time.

We have already discussed the results of training a response repertory that is too narrow. A program can also be built to teach the student to generalize. Programs are not the only teaching procedures that can suppress originality and spontaneity in students when these procedures are employed at their worst. There is no evidence to suggest that spontaneity is decreased through

programs per se, but a great deal of evidence has been accumulated over the agony of years to indicate that spontaneity and curiosity, initiative and resourcefulness have carefully been conditioned out of students who have gone through the traditional primary and secondary education. The impressive thing about this accomplishment is that it was done without the aid of automation.

What does programmed instruction do to the originality of a student's approach to learning? Does he become overly dependent upon this type of instruction? Skinner reports that students tend to be more spontaneous in conventional teaching situations, such as discussion groups, when they have the confidence of well-mastered knowledge (Skinner, 1961). Granted that opinions must at present be based on impression, nothing suggests over-dependence on programmed instruction. Students seem to be more highly motivated and certainly better prepared as a consequence of learning through this kind of instruction. This is, however, a legitimate question for empirical research.

Socio-political Implications

Will programmed instruction lead to totalitarianism? If a program is designed to teach Communism or Fascism or Democracy then that is what the program will teach, just as a book will do. The machine may do a more or less effective job, just as different teachers may do more or less effective jobs in propagandizing students. There is nothing about a machine or a program, just as there is nothing about a book or a teacher, that forces the

student to adopt one point of view over competing or alternative points of view. In fact, it should be possible to teach through programming in such a way that a more, rather than a less, critical attitude is developed. It may be possible to teach principles of Democracy, not through a Pollyannish emphasis on the good things of our culture and a studied disregard of its inconsistencies, but rather through teaching an analytic attitude. If students were taught to ask the same searching questions of competing systems that they asked of ours, they would inevitably reject inferior systems.

In brief, the technique of programmed instruction constitutes a kind of discrimination training on the part of the teacher which makes use of every effective stimulus-response-reinforcement contingency. The effect of this training is more powerful than that of conventional techniques of instruction. A more effective technique is bound to produce sharper discriminations. New behavioral repertories will be expanded and more firmly established than would be possible with older, less precisely controlled conditions of learning. The net effect cannot help but be advantageous, both to the students and to the society they ultimately will create. If certain institutions are altered as a consequence of the behavior of people who are more adequately prepared to deal with their environment, then those are institutions that should be altered in the interests of the survival of this civilization. What we hope to accomplish with more effective techniques of training is the introduction of more effective modes of action. Where those more effective modes of action lead to changes, we should welcome such changes.

The Technical Questions

We may return to certain technical questions of general concern to the prospective user of programs. One of these is: How does one write a good program? The sixth chapter of this book discusses in some detail techniques of presentation of material and devices that are employed in program writing. Several excellent papers have been published on the technique of program writing. The final answer, however, to the question of how one goes about writing a good program is that one writes a poor one. One must write a poor program, try it, and learn what makes it a poor program. One then revises it so that it is no longer a poor program. Unfortunately, no one can tell another person how to write. Not everyone can become a good novelist; not everyone can write a good book; and not everyone can write a good program. We can only say that the person who writes a good program must know his subject matter. He knows it in the sense that he uses it in novel situations; he can discuss it imaginatively. He is interested in his subject matter, he enjoys it, he finds it personally stimulating and rewarding. He is close enough to the student to be able to see the student learn. He is not so bound by traditional concepts of human behavior that he cannot learn by observing how conditioning or learning really occurs. He is able to learn what reinforcement looks like in actual practice. He is able to employ or has already employed reinforcement. He has considerable patience, because writing a program takes time. He knows what the student should know because he lets the student tell him. He does not have preconceived notions of how the stu-

dent should learn or what he should know. The good programmer follows the dictum: The subject is always right.

Some final questions we might ask are: What should be programmed? What are the objectives of programming? What is desirable behavior? Answers to these questions can be given in several different contexts. In the practical context, the answer is that programmed instruction will seek to teach those things which can most effectively be taught by this method. We should recognize that teaching machines and programs will not solve all our educational problems.

Resistance to Programmed Instruction

As is always true of the introduction of any new concept or technique, several objections have been raised to programmed instruction. Not all such objections are meaningless, and it is incumbent upon us to answer some of them. What is the nature of these objections? Where do they originate? Objections are introduced at all levels. Some persons react emotionally to the name "teaching machine." The choice of this term may be unfortunate, but the reaction to it is even more so. As is by now apparent to the reader, there is little of naive mechanism in programmed instruction. The term, teaching machine, suggests to many persons an infernal device leading us irrevocably into some grisly totalitarian society (Orwell, 1949; Huxley, 1932). Such pessimism is groundless. But the technique of programmed instruction is equally unlikely to deliver us into the kind of utopia envisioned by

Skinner (Skinner, 1948). It is simply a more efficient mode of instruction. Some of the reaction to it is the contemporary counterpart of the resistance that medieval illuminators must have shown to the printing press, that soulless engine of mass-produced learning.

Programmed instruction is no more likely to "brainwash" students than is the biased lecture of a gifted platform artist. If one objects to this technique of instruction on the basis that it teaches more efficiently, one basically objects to teaching. Such objections do exist. It is hard to recognize, but many persons prefer a state of ignorance, both for themselves and for others. Such persons must react violently against any threat to entrenched ignorance. Although they may make vast changes in our educational system and concepts, schools will remain schools. The invention of the printing press caused drastic changes in the structure of education, but certain features of the educational system survived. The lecture, as it was originally read, was the reading to students of material to which they did not have general access. With the invention of the printed book, students have acquired access to this material. In a real sense, the present technological advance is directly analogous to the introduction of the printing press in that material that was formerly inaccessible to the student now becomes accessible. In both cases, material that had to be taught by the teacher is relegated to the supplemental technique, on one hand, the textbook, on the other, the program. The textbook did not result in the technological unemployment of the lecturer. The development of the teaching machine or of the programmed textbook will not turn the teacher into a part-time mechanic. Komoski's comment is perhaps the last word in this regard: "Any teacher

who can be replaced by a machine deserves to be replaced" (Komoski, 1961).

Other objections are voiced. For example, one public official responsible for educational policies voiced the fear that these devices "threaten to destroy diversity and place pupils in molds." It is sad that such misconceptions exist. As is evident from the present analysis, programmed instruction holds forth the first ray of hope that the abysmal rigidities of a curriculum-bound educational system might yield up their victims for some measure of intellectual resuscitation. For the sake of clarification, let it be recorded that the present educational system is most marked by conformity and regimentation. Acceleration of bright students is enough of a rarity to represent a point of pride for the exceptional school that can accomplish it to any extent. If allowed to, programmed instruction will have the opposite effect of placing students in identical molds. The only equalizing effect that programmed instruction presumably will have is the education of everyone to a basic minimal level. If this is what is feared, and if conventional teaching does not accomplish this, then we must ask: Is this *bad*?

Popular writers have played upon the theme that teachers will be relegated to the status of glorified mechanics by machines. That this is distorted sensationalism is readily apparent. The only threat to the teacher is that the occasional second-rate one may someday face a class fully as well informed concerning basic facts as himself. Programmed instruction may make teaching more difficult in one sense.

The teacher cannot go to a standard textbook to develop his lecture. He is forced to go to primary sources. This consultation of primary sources keeps the teacher

up-to-date and presumably makes the material he presents more alive for the student. It is unfortunately true that such keeping up-to-date will be an aversive task for many teachers, but it is one for which they are long overdue. Going to primary sources reinstates the validity of the lecture in the old sense of the word. The main value of the lecture is in making available to the students material to which they otherwise have limited access. From personal experience, we know that it is desirable to send students, particularly in research disciplines, to original journal articles. It is for this reason that so many collections of readings abound in the various disciplines. Unfortunately, however, it is seldom true that a set of readings keyed to a particular lecturer's presentation or emphasis is appropriate to another. If the lecturer were free to devote his time to presenting those basic materials to which it is not feasible to send students, rather than devote his time to recapitulating material that exists in conventional textbooks, a more effective job could be done. As stated before, the Renaissance lecturer served precisely this function. He presented to the students material that was otherwise not available to them. His was a legitimate function of education, a far cry from the deteriorated status of the lecturer in an introductory course in psychology or sociology where the lecturer often engages in the redundant channeling into the auditory modality material that is otherwise available to the visual.

Objections abound on the theoretical level. Certain persons have decried the invasion of stimulus-response psychology into the classroom (Snygg, 1961). Such objections are of little relevance to the practical use of programmed instruction for they represent only the familiar internecine epistomological strife between Gestalt and

behavioristic psychologies. The objections have little real relevance to education, and one wonders whether they would ever have been raised had the opposing school been able to develop successful techniques of control.

The teacher must beware of the more involved theoretical arguments bearing upon programmed instruction; they often generate more heat than light. The unique contribution of the learning psychologist is in presenting what he knows of basic processes, stripped of theoretical jargon and interpretations, and in showing the techniques that have been used to study behavior. He should do this in such a way that the teacher can apply these same techniques to his own problems. There is a difference of emphasis in the approach the experimental psychologist takes to the learning process and the approach the teacher takes. The experimental psychologist is concerned primarily with the fundamental processes underlying behavior, whereas the teacher is concerned with the end product of those processes. The teacher, however, can learn much concerning the kinds of variables which he can manipulate to increase the efficiency of teaching. The greatest danger in the teacher's coming into contact with the experimental psychologist is in the difficulty of discriminating between what is descriptive fact and what is theoretical fiction. No psychologist who has personally conditioned an animal would speak of children's learning as the pouring-in of information that is to be stored at random in a child's brain. Nor would any experimental psychologist who was concerned with the descriptive analysis of behavior be likely to conclude that machine teaching is based upon a hierarchical storage scheme where instruction supplied the student

with some kind of routing system. Such verbalisms may sound appealing to the uninformed, but not only do they not add anything to our understanding of the way in which learning occurs, but they confuse the issue by imposing a computing machine analogy that is pure fiction as a description of the actual function of the human nervous system. How can one teach children more effectively with concepts such as storage banks and routing systems? Theory has a legitimate place in science. One function of theory is to summarize efficiently and economically the statements of fact that have been accumulated through systematic empirical study. A second function is to indicate the most fruitful direction of the next experiment. Verbalisms that have the surface appearance of theory but serve neither of these roles are illegitimate in the extreme.

Practical Considerations

More legitimate objections have to do with the practical realities. Do machines really teach? If the answer to this turns out to be no, then it is the ultimate question. The only way to answer the question is to try. We know that machines, programs, do teach some things quite well. How generally effective are they? Should we proceed with a technique when we don't as yet know its limitations? Unfortunately, the limitations have to be found empirically. Mistakes will be made. Even if it should develop that programmed instruction has a very limited usefulness, much has been and will be learned in the exploration of its limits. The fact that public concern and

research efforts are directed toward problems of learning represents pure gain.

What about the desirability of linear programs versus branching programs? First, there is no reason that these alternatives should be regarded as exhaustive of all the possible forms of programs. Rigid clinging to one school or another befits the behavior of disciples of a cult more than it does the behavior of scientists. The real challenge now is to develop the program of tomorrow, not to justify the primitive state of knowledge of yesterday—regardless of one's persuasion on this or any other issue. Different forms of programs are likely to be effective when fitted to the task. Several writers have referred to the response mode required of programs as crucial in their evaluation. I would argue that programs should be fitted to the response mode. A program that is designed to differentiate a new class of behavior would work somewhat differently than one designed to bring existing classes of behaviors under the discriminative control of the environment. An effective program for teaching spelling might look different from a program for teaching accounting. We most likely shall reject both existing forms of programs when we have better learned how to program. The present arguments have been intended to evaluate the existing alternatives but they should not be taken to defend any status quo.

How might we go about developing new programs? The first step is to understand the properties of those we have. It has been possible to formulate rigorous analyses of grammar in terms of the parts of speech. It must be feasible to formulate analyses of programs in terms of their formal properties and functions. We have some linguistic categories to begin with. We refer to S^D and S^Δ

occasioners in discrimination training, and we know something of their function and constitution. Density functions are a step in the formal descriptions of programs, as is the ruleg system. We can categorize programs which differentiate and which discriminate. We know something about how to teach generalization. One can define a program as an inductive or as a deductive program. We can specify the type and number of prompts in a program or subprogram. The response modality must enter somewhere. In short, there are any number of ways to start the analysis of the formal properties of programs. Most of them probably will not bear fruit. Some of them most assuredly will lead us in the direction of developing more effective programs.

What about machines? We have raised some objections to the premature freezing of technique to accommodate a particular device. There is a tendency among educators to commit themselves to inferior devices. Programmed instruction promises to be a financially rewarding development. As such, it is a source of temptation for those who seek to exploit. Not only machines but also programs may be inferior. It is not true that *anything* teaches better if it is put into a machine.

Another danger with machines is that certain manufacturers are exploiting the present interest in automated instruction to market devices that do automate certain aspects of classroom conduct but do not incorporate the defining features of *teaching machines*. One finds, for example, proposals for mass testing that utilize elaborate scoring devices. Some of these devices make no provision for scoring individual students upon their performance; almost all require that the class be paced to the speed of the slowest student, and some even claim that the

device does away with the need for a program! These "systems" are uniformly expensive. Our only concern with such devices is to emphasize that such contrivances do *not* meet the criteria defining programmed instruction.

The greatest threat is that programmed instruction will be oversold. Great claims have been made for the technique. Perhaps the technique cannot meet all claims that are made for it. No true panacea has been found. Those who expect too much are bound for disappointment. The disappointment might become a reaction against the technique that could cause its total rejection in the long run. This would be unfortunate. It is the reason why caution must be used in the promotion of machines and programs.

The Potentialities of Programmed Instruction

What might we legitimately expect of this new technique? We cannot expect that it will make learning painless. We shall never, science fiction notwithstanding, be able to put our heads into a gadget that will instantaneously impart to us the wisdom of the ages or fluency in a new language. Learning is behavior, and behavior is work. Programmed instruction will not solve all the educational ills of our society. It will not singlehandedly win the ideological struggle against our enemies. They can use it too. Much more is involved here.

If the student-tutor relationship is an effective teaching situation, then insofar as programmed instruction approximates that relationship, it, too, will be an effec-

tive teaching situation. If the best tutors program their knowledge into good programs, the educational gain could be tremendous. Imagine what it could mean to the student in the rural high school or small college to be able to learn in such a setting from the Einsteins, the Pasteurs, the Shakespeares, and even the Toscaninis of our civilization.

Deficits in preparation could be more easily remedied. The child who missed fractions because he had the measles that week need no longer ever after lack "aptitude for mathematics." The graduate student who lacks certain preparation need not defer his graduate career while he takes time off to make up those lacks. Such supplementary education could be undertaken concurrently with the formal courses of study. More time might be released for the pursuit of independent research at all levels. Students could be more quickly brought to a point where independent research would be meaningful. The time required for mastering the specialties might be reduced. The impersonal relationship between student and machine or student and program should result in less negative reaction on the part of the student. The program does not threaten or punish. It can only reinforce. An important advantage lies in the therapeutic effect on education resulting from such a change in technique. Many learning difficulties are emotional rather than intellectual. Reducing the incidence of these could have incalculable advantages.

Only one thing is certain. No matter what the potentialities of this technique may be, it is a tool. It will be only as good or as bad as the uses to which it is put. All the argument favoring programmed instruction and all the polemics damning it are but breezes in the trees.

The justification of this development, the justification of all conceivable applications of laboratory technique to practical use in the affairs of men, lies in their future success. Such matters cannot be settled by debate. We must try and see.

Appendix

The following is a statement on self-instructional materials and devices released by a joint committee of the American Educational Research Association, the American Psychological Association, and the Department of Audio-visual Instruction of the National Education Association.

> The use of self-instructional programmed learning materials in teaching machines and similar devices represents a potential contribution of great importance to American education. But this contribution can best be realized only if users have information with which to evaluate self-instructional materials. Accordingly, the following interim guide-lines have been prepared.
>
> 1. Teaching machines do not, in themselves, teach. Rather, the teaching is done by a program of instructional materials presented by the teaching machine. Any evaluation of a teaching machine thus requires an assessment of the availability and quality of programs for each type of machine, as well as its mechanical dependability.
>
> 2. A variety of programmed materials is becoming available, but not all programs will fit all machines. Thus

only those programs compatible with a particular machine can be considered as available for use with it. A list of commercially available programs and devices can be obtained by sending a request to the Department of Audio-visual Instruction, National Education Association.

3. In evaluating the specific content which a self-instructional program purports to teach, the program can be examined to determine what the student is required to do and whether the student's responses reflect the kind of competence which the educator wishes to achieve. Like other educational materials, programs labeled with the name of a particular subject matter vary widely with respect to content and instructional objectives.

4. Just any set of question and answer material does not constitute a self-instructional program. One major type of self-instructional material proceeds by small steps requiring frequent student responses. These steps can be examined to see if they embody a careful, logical progression of the subject matter. Items in such a program are designed so that the student will respond to the critical aspects of each item or will perform the important operation which that item was meant to teach. Furthermore such programs generally provide a wide range of examples illustrating each principle or concept.

5. Self-instructional materials are designed to adapt to individual differences by allowing each student to proceed at his own rate. Some types of self-instructional materials further adapt by "branching" to alternate materials. For this purpose, questions are designed to diagnose the student's needs, and to provide alternate material suited to these needs. The material is designed so that the choice of answer to a particular question determines which items will be presented next. Incorrect answers take the student to items containing information

designed to correct the error before continuing through the sequence.

6. An important feature of almost all self-instructional materials is that a record of the student's responses provides a basis for revising the program. The prospective purchaser should ask about the extent to which revision has been based on student response and how the preliminary tryout was conducted.

7. The effectiveness of a self-instructional program can be assessed by finding out what students actually learn and remember from the program. The prospective purchaser should find out whether such data are available and for what kinds of students and under what conditions the data were obtained.

8. Active experimentation with self-instructional materials and devices in school systems is to be encouraged prior to large-scale adoption.

References

Amsel, A. (1960) Error responses and reinforcement schedules in self-instructional devices. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning.* Washington, D.C.: Natl. Educ. Ass., Dep. of Audio-visual Instruction.

Bridgman, P. W. (1927) *The logic of modern physics.* New York: Macmillan.

Bullock, D. H., & Maline, D. B. (1958) Diagnosing child behavior. *The Training School Bull.,* 55, pp. 47–53.

Carr, W. J. (1960) A functional analysis of self-instructional devices. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning.* Washington, D.C.: Natl. Educ. Ass., Dep. of Audio-visual Instruction.

Coulson, J. E., & Silberman, H. F. (1960) Results of an initial experiment in automated teaching. *J. educ. Psychol., 51,* pp. 135–143.

Crowder, N. (1959) Automatic tutoring by means of intrinsic programming. In E. H. Galanter (Ed.), *Automatic teaching: the state of the art.* New York: John Wiley and Sons.

Edwards, W. D. (May 1956) Skinner's teaching machines. Unpublished laboratory note. Lowry Air Force Base, Colorado: *AFPTRC,* Maintenance Lab.

Eigen, L. D., & Komoski, P. K. (1960) Research summary #1 of the Collegiate School automated teaching project. New York.

Estes, W. K. (1944) An experimental study of punishment. *Psychol. Monogr., 57,* No. 263.

Estes, W. K. (1950) Towards a statistical theory of learning. *Psychol. Rev., 57,* pp. 94–107.

Estes, W. K. (1955) Statistical theory of spontaneous recovery and regression. *Psychol. Rev., 62,* pp. 145–154.

Estes, W. K., & Straughan, J. H. (1954) Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol., 47,* pp. 225–234.

Estes, W. K., *et al.* (1954) *Modern learning theory.* New York: Appleton-Century-Crofts.

Evans, J. L., Homme, L. E., & Glaser, R. (1960) The ruleg (rule-example) system for the construction of learning programs. A report prepared under the Cooperative Research Program of the U.S. Office of Educ., Univer. of Pittsburgh, Dept. of Psychol.

Ferster, C. B. (1957) Withdrawal of positive reinforcement as punishment. *Science, 126,* p. 509.

Ferster, C. B., & Skinner, B. F. (1957) *Schedules of reinforcement.* New York: Appleton-Century-Crofts.

Foltz, C. I. (1961) *The world of teaching machines.* Washington, D.C.: Electronic Teaching Laboratories, Teaching Res. and Technol. Div.

Frank, P. (1957) *Philosophy of science, the link between science and philosophy.* New York: Prentice-Hall.

Galanter, E. H. (1959) The ideal teacher. In E. H. Ga-

lanter (Ed.), *Automatic teaching: the state of the art*. New York: John Wiley and Sons.

Gibson, J. J. (1960) The concept of the stimulus in psychology. *Amer. Psychologist, 15*, pp. 694–703.

Gilbert, T. F. (1960) On the relevance of laboratory investigation of learning to self-instructional programming. In A. A. Lumsdaine & R. Glaser, (Eds.), *Teaching machines and programmed learning, a source book*. Washington, D.C.: Natl. Educ. Ass., Dep. of Audio-visual Instruction.

Green, E. J. (1955) Concept formation, a problem in human operant conditioning. *J. exp. Psychol., 49*, pp. 175–180.

Green, E. J. (1956) Stimulus variability and operant discrimination in human subjects. *Amer. J. Psychol., 69*, pp. 269–273.

Green, E. J. (1958) A simplified model for stimulus discrimination. *Psychol. Rev., 65*, pp. 56–63.

Green, E. J. (1961) A technique for determining the densities of programs. *AID, 1*, No. 3, p. 27.

Green, E. J. (1961a) Error rate and response variability in programmed instruction. Unpublished manuscript.

Green, E. J., & Meining, F. R. (1957) Rate response probability in discrimination learning. *Science, 126*, pp. 1339–1340.

Green, E. J., Sanders, R. M., & Squier, R. W., Jr. (1959) Schedules of reinforcement and discrimination learning. *J. exp. anal. Behav., 2*, pp. 293–299.

Greenspoon, J. (1955) The reinforcing effects of two spoken words on the frequency of two responses. *Amer. J. Psychol., 68*, pp. 409–416.

Guthrie, E. R. (1935) *The psychology of learning.* New York: Harper & Bros.

Hiebsch, H., *et al.* (1961) *Soviet psychology—a symposium.* New York: Philosophical Lib.

Holland, J. G. (1958) Human vigilance. *Science, 128,* pp. 61–63.

Holland, J. G. (1959) Teaching psychology by a teaching machine program. Paper read at the Amer. Psychol. Ass., Cincinnati.

Holland, J. G. (1960) Teaching machines: an application of principles from the laboratory. *J. exp. anal. Behav., 3,* pp. 275–287.

Holland, J. G., & Skinner, B. F. (1961) *The analysis of behavior.* New York: McGraw-Hill.

Homme, L. E., & Glaser, R. (1960) Problems in programming verbal learning sequences. In A. A. Lumsdaine, & R. Glaser (Eds.), *Teaching machines and programmed learning.* Washington, D.C.: Natl. Educ. Ass., Dep. of Audio-visual Instruction.

Hull, C. L. (1933) *Hypnosis and suggestibility: an experimental approach.* New York: Appleton-Century-Crofts.

Hull, C. L. (1943) *Principles of behavior.* New York: Appleton-Century-Crofts.

Hull, C. L. (1951) *Essentials of behavior.* New Haven, Conn.: Yale Univer. Press.

Hull, C. L. (1952) *A behavior system: an introduction to behavior theory concerning the individual organism,* New Haven, Conn.: Yale Univer. Press.

Huxley, A. L. (1932) *Brave new world.* New York: Harper & Bros.

Irwin, O. C. (1947) Infant speech: consonantal sounds

according to place of articulation. *J. speech Disorders, 12,* pp. 397–401.

Komoski, P. K. (1961) Quotations in *Time, 77,* p. 38.

McCrary, J. W. (1956) Personal communication.

Markle, S. M., Eigen, L. D., & Komoski, P. K. (1961) *A programed primer on programing.* New York: Center for Programed Instruction.

Mechner, F. (1961) *A teaching machine program on programmed instruction.* New York: Basic Systems, Inc.

Mellen, I. (1936) Teaching and educational inventions. *J. exp. Educ., 4,* pp. 291–300.

Meyer (Markle), S. R. (1960) Report on the initial test of a junior high school vocabulary program. In A. A. Lumsdaine & R. Glaser (Eds.), *Teaching machines and programmed learning.* Washington, D.C.: Natl. Educ. Ass., Dep. of Audio-visual Instruction.

Moore, O. K., & Andrews, A. R. (1959) *Early reading and writing.* A film report. Little Harbor, Guilford, Conn.: Basic Educ., Inc.

Mowrer, O. H., & Jones, H. M. (1945) Habit strength as a function of the pattern of reinforcement. *J. exp. Psychol., 35,* pp. 293–311.

Orwell, G. (1949) *1984.* New York: Harcourt, Brace & World.

Pask, G. (1958) Electronic keyboard teaching machines. *Education and Commerce.*

Pavlov, I. P. (1927) *Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex.* London: Oxford Univer. Press.

Rigney, J. W., & Fry, E. B. (1961) Current teaching machine programs and programming techniques. *AV Communication Review, 9,* No. 3

Skinner, B. F. (1935) The generic nature of the concepts of stimulus and response. *J. gen. Psychol., 12,* pp. 40–65.

Skinner, B. F. (1938) *The behavior of organisms: an experimental analysis.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1948) *Walden two.* New York: Macmillan.

Skinner, B. F. (1948a) Superstition in the pigeon. *J. exp. Psychol., 38,* pp. 168–172.

Skinner, B. F. (1950) Are theories of learning necessary? *Psychol. Rev., 57,* pp. 193–216.

Skinner, B. F. (1954) The science of learning and the art of teaching. *Current Trends in Psychology and the behavioral Sciences.* Pittsburgh.

Skinner, B. F. (Winter 1955–1956) Freedom and the control of men. *The American Scholar.*

Skinner, B. F. (1957) *Verbal behavior.* New York: Appleton-Century-Crofts.

Skinner, B. F. (1958) Reinforcement today. *Amer. Psychologist, 13,* pp. 94–99.

Skinner, B. F. (1961) Reply to question in inside opinion. *Programed Instruction, 1,* p. 1.

Snygg, D. (1961) Address to Div. of Audio-visual Instruction at Natl. Educ. Ass. Convention, 1960. Reported in *Audio-visual Instruction, 6,* p. 276.

Thorndike, E. L. (1913) The psychology of learning. (Educational Psychol. 11), New York: Teachers College, Columbia Univer.

Thorndike, E. L. (1932) Reward and punishment in animal learning. *Comp. Psychol. Monogr., 8,* No. 39.

Tolman, E. C. (1932) *Purposive behavior in animals and men.* New York: Appleton-Century-Crofts.

Turing, A. M. (1950) Computing machines and intelligence. *Mind, 59,* pp. 433–460.

Verplanck, W. S. (1955) The control of the context of conversation: reinforcement of statements of opinion. *J. abnorm. soc. Psychol., 51,* pp. 668–676.

Verplanck, W. S. (1956) The operant conditioning of human motor behavior. *Psychol. Bull., 53,* pp. 70–83.

Verplanck, W. S., & Hayes, J. R. (1953) Eating and drinking as a function of maintenance schedule. *J. comp. physiol. Psychol., 46,* pp. 327–333.

Wallach, H., & Henle, M. (1941) An experimental analysis of the law of effect. *J. exp. Psychol., 28,* pp. 340–349.

Watson, J. B. (1913) Psychology as the behaviorist views it. *Psychol. Rev., 20,* pp. 158–177.

Wiener, N. (1948) *Cybernetics: or control and communication in the animal and the machine.* New York: John Wiley and Sons.

Zeaman, D. (1959) Skinner's theory of teaching machines. In E. H. Galanter (Ed.), *Automatic teaching: the state of the art.* New York: John Wiley and Sons.

Index